Chapter 705

Group-Sequential Superiority by a Margin T-Tests for Two Means

The corresponding sample size procedure, found in PASS Sample Size software, is Group-Sequential Superiority by a Margin T-Tests for Two Means (Simulation).

Introduction

This procedure is used to test superiority by a margin for the difference of two means in stages (sometimes called looks or interim analyses). Unless the stage boundaries are entered directly, the stage boundaries are defined using a specified spending function. Futility boundaries can be binding or non-binding. Futility boundaries are specified through a beta-spending function.

Sample size re-estimation, based on current-stage sample sizes and parameter estimates, may also be obtained in this procedure.

The spending functions available in this procedure are the O'Brien-Fleming analog, the Pocock analog, the Hwang-Shih-DeCani gamma family, and the power family.

At each stage, the current and future boundaries are calculated based on the accumulated information proportion. Conditional and predictive power for future stages is also given. Sample size adjustments (re-estimation), based on current estimates of group standard deviations, can be given.

A group-sequential boundary and analysis graph is produced in this procedure.
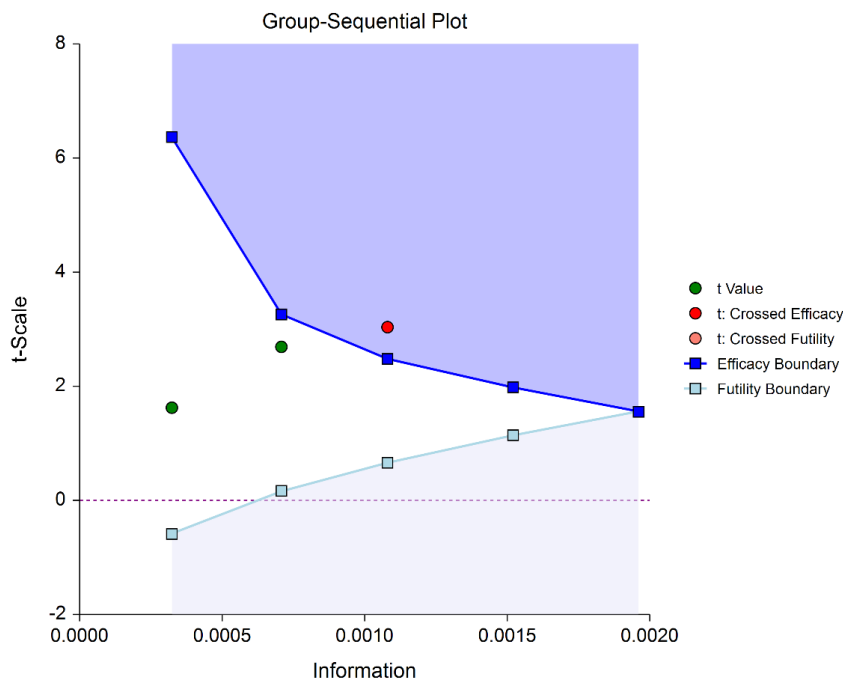
At each stage, stage-adjusted difference estimates, confidence intervals, and p-values are available.

The probabilities of crossing future boundaries may also be assessed, using simulation.

The format of the data for use in this procedure is three columns: one column for the response values, one column defining the two groups, and a third column defining the stage.

Outline of a Group-Sequential Study

There are three basic phases of a group-sequential (interim analysis) study:

- Design

- Group-Sequential Analysis

- Reporting

Design Phase – Determine the Number of Subjects

To begin the group-sequential testing process, an initial calculation should be made to determine the sample size and target information if the final stage is reached (maximum information). The sample size calculation requires the specification of the following:

- Alpha

- Power

- Test Direction

- Types of boundaries (efficacy, binding futility, non-binding futility)

- Maximum number of stages

- Proportion of maximum information at each stage

- Spending functions

- Within group standard deviations

- Assumed difference in means

The design phase calculation may be done in the **PASS** sample size software program. **PASS** software permits the user to easily try a range of standard deviations and mean differences, as these values are typically not known in advance.

The resulting sample size of the sample size calculation also permits the calculation of the maximum information, which is the total information of the study if the final stage is reached (for calculation details, see the Information section later in this chapter).

Based on the maximum information, the target information and target sample size of each stage may be calculated. In particular, this permits the user to have a target sample size for the first stage.

Although it is likely to change over the course of the group-sequential analysis, a design group-sequential boundary plot can be a useful visual representation of the design:

Group-Sequential Plot
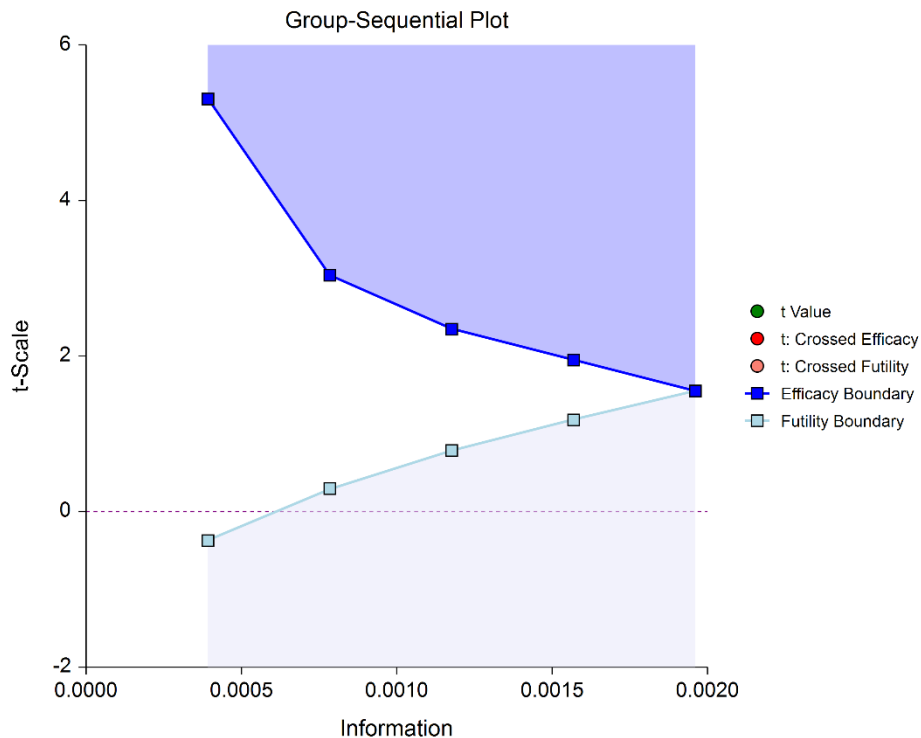
Group-Sequential Analysis Phase

A group sequential analysis consists of a series of stages where a decision to stop or continue is made at each stage.

First Interim Stage

The design phase gives the target number of subjects for the first stage. The study begins, and response data is collected for subjects, moving toward the first-stage target number of subjects, until a decision to perform an analysis on the existing data is made. The analysis at this point is called the first stage.

Unless the number of subjects at the first stage matches the design target for the first stage, the calculated information at the first stage will not exactly match the design information for the first stage. Further, the sample standard deviations will rarely, if ever, match the standard deviations used in the calculation of the information at the design stage, and thus the calculated information at the first stage will differ from the design information. Generally, the calculated information will not differ too greatly from the design information, but regardless, spending function group-sequential analysis is well-suited to make appropriate adjustments for any differences.

The first stage information is divided by the maximum information to obtain the stage one information proportion (or information fraction). This information proportion is used in conjunction with the spending function(s) to determine the alpha and/or beta spent at that stage. In turn, stage one boundaries, corresponding to the information proportion, are calculated.

A *t*-statistic is calculated from the raw mean difference and sample standard deviations. The stage one *t*-statistic is compared to each of the stage one boundaries. Typically, if one of the boundaries is crossed, the study is stopped (non-binding futility boundaries may be an exception).
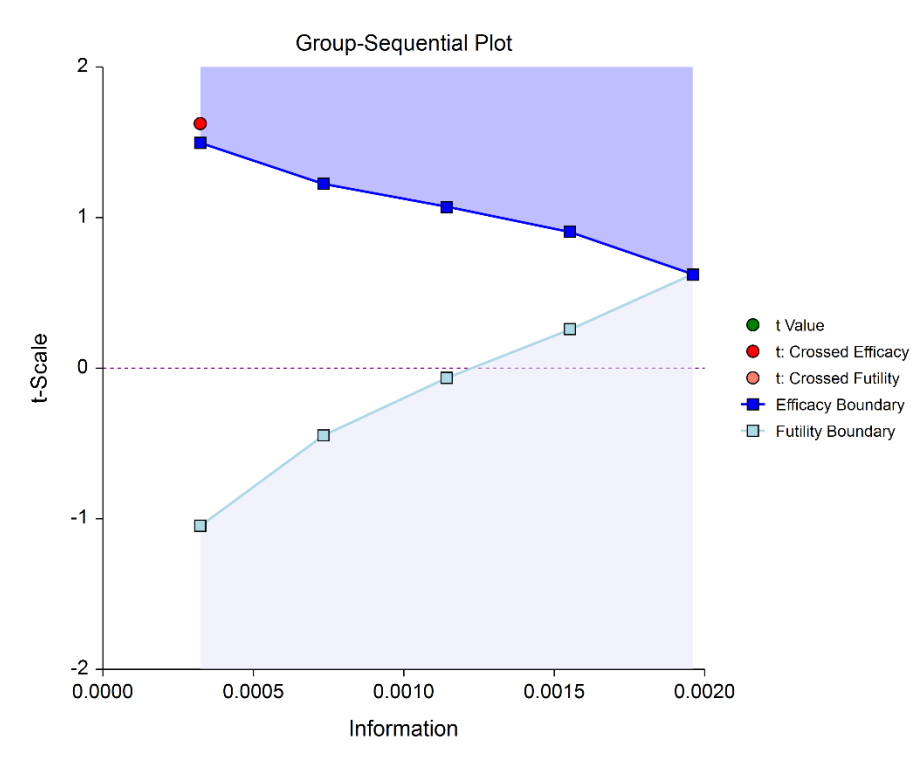
If none of the boundaries are crossed the study continues to the next stage.
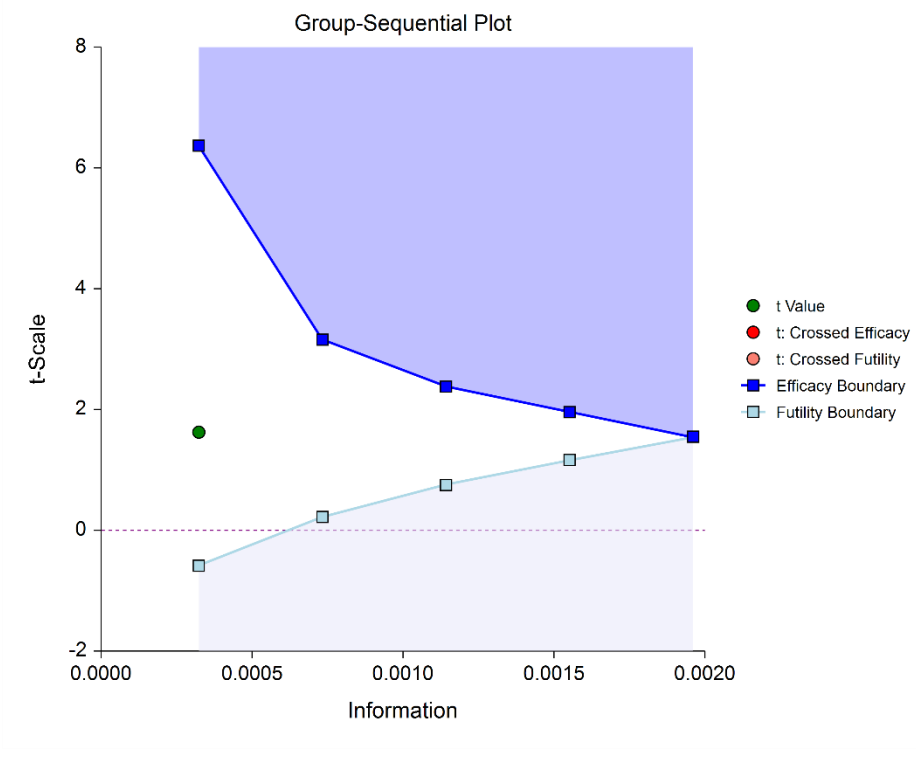
If none of the boundaries are crossed it may also be useful to examine the conditional power or stopping probabilities of future stages. Conditional power and stopping probabilities are based on the user-specified supposed true difference.

Second and other interim stages (if reached)

Since the first stage information proportion is not equal to the design information proportion, a designation must be made at this point as to the target information of the second stage. Two options are available in this procedure.

One option is to target the information proportion of the original design. For example, if the original design proportions of a four-stage design are 0.25, 0.50, 0.75, 1.0, and the stage one observed proportion is 0.22, the researcher might still opt to target 0.50 for the second stage, even though that now requires an additional information accumulation of 0.28 (proportion). The third and fourth stage targets would also remain 0.75 and 1.0.

A second option is to adjust the target information proportionally to the remaining proportions. For this option, if the design proportions are 0.25, 0.50, 0.75, 1.0, and 0.22 is observed, the remaining 0.78 is distributed proportionally to the remaining stages. In this example, the remaining target proportions become 0.48, 0.74, 1.0.

For either option, once the target information is determined for the next stage, revised target sample sizes are given, and the study continues until the decision is made to perform the next interim analysis on the cumulative response data. In the same manner as the first stage, the current stage information proportion is used with the spending function to determine alpha and/or beta spent at the current stage. The current stage boundaries are then computed. The *t*-statistic is calculated and compared to the boundaries, and a decision is made to stop or continue.

If a boundary is crossed, the study is typically stopped.
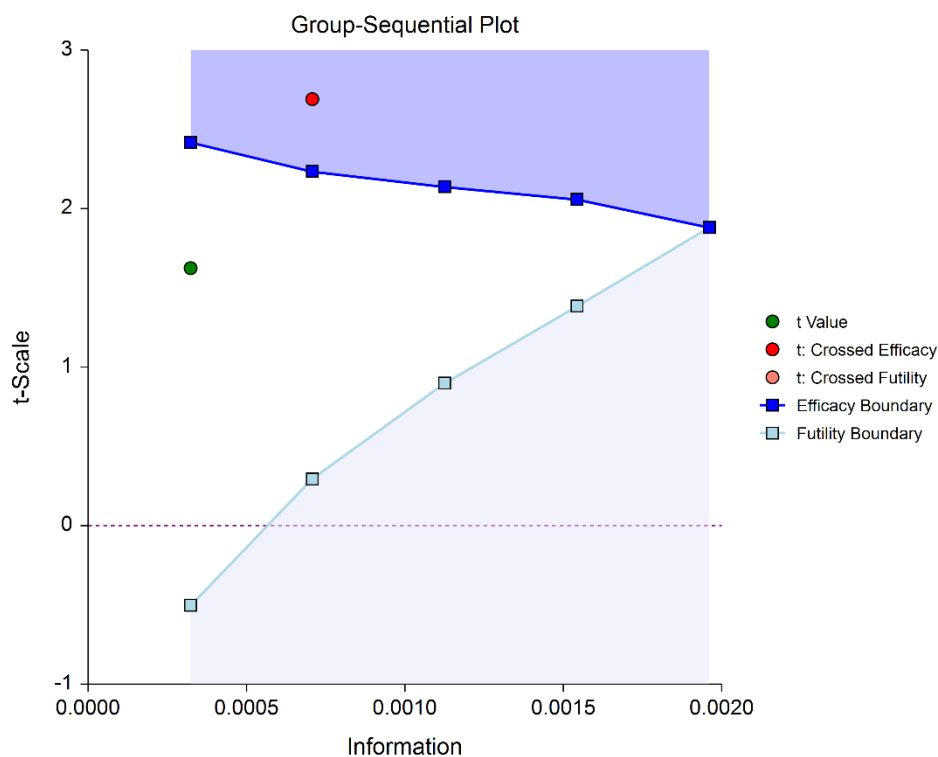
If none of the boundaries are crossed the study continues to the next stage.
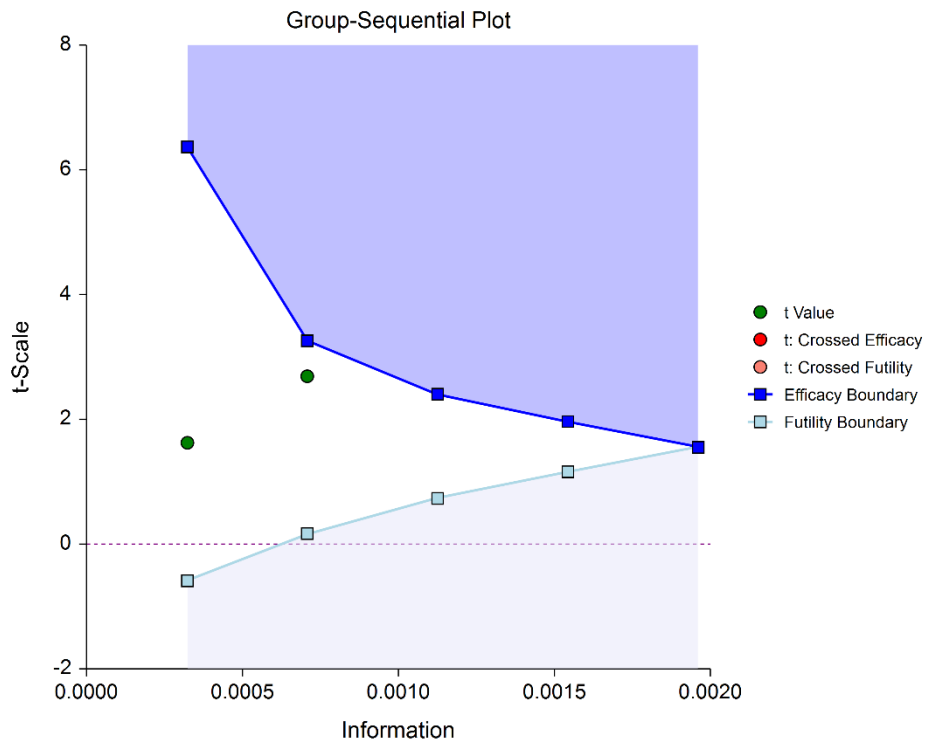
Once again, if no boundary is crossed, conditional power and stopping probabilities may be considered based on a choice of a supposed true difference.

The study continues from stage to stage until the study is stopped for the crossing of a boundary, or until the final stage is reached.

Final Stage (if reached)

The final stage (if reached) is similar to all the interim stages, with a couple of exceptions. For all interim analyses the decision is made whether to stop for the crossing of a boundary, or to continue to the next stage. At the final stage, only the decision of efficacy or futility can be made.

Another intricacy of the final stage that does not apply to the interim stages is the calculation of the maximum information. At the final stage, the current information must become the maximum information, since the spending functions require that the proportion of information at the final look must be 1.0. If the current information at the final stage is less than the design maximum information, the scenario is sometimes described as *under-running*. Similarly, if the current information at the final stage is greater than the design maximum information, the result may be termed *over-running*.

For both under-running and over-running, the mechanism for adjustment is the same, and is described in the Technical Details section, under Information and Total Information.

Aside from these two exceptions, the final stage analysis is made in the same way that interim analyses were made. The remaining alpha and beta to be spent are used to calculate the final stage boundaries. If the test is a one-sided test, then the final stage boundary is a single value. The final stage *t*-statistic is computed from the sample means and standard deviations of the complete data from each group. The *t*-statistic is compared to the boundary and a decision of efficacy or futility is made.

Reporting Phase

Once a group-sequential boundary is crossed and the decision is made to stop, there remains the need to properly summarize and communicate the study results. Some or all of the following may be reported:

- Boundary plot showing the crossed boundary
- Adjusted confidence interval and estimate of the mean difference
- Sample size used

Boundary plot showing the crossed boundary

The boundary plot gives an appropriate visual summary of the process leading to the reported decision of the study.
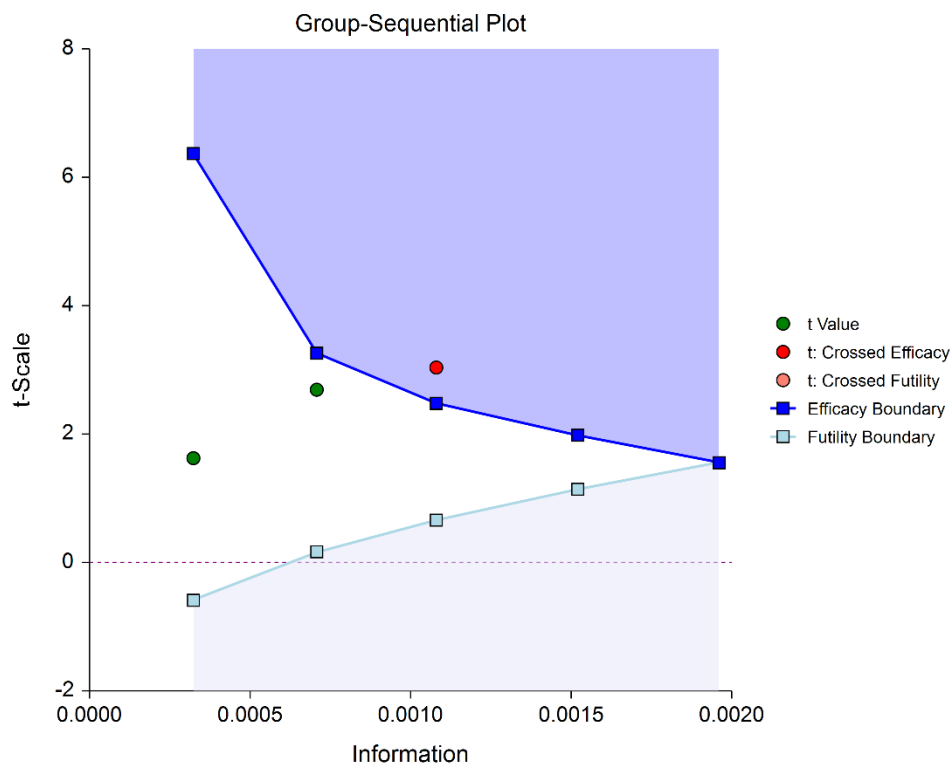
Adjusted confidence interval and estimate of the mean difference

Due to the bias that is introduced in the group-sequential analysis process, the raw data confidence interval of the difference in means should not be used. An adjusted confidence interval should be used instead. See the Adjusted Confidence Interval topic of the Technical Details section for details. The mid-point of the adjusted confidence interval may be useful as a point estimate of the difference in means. Further, the confidence level at which the adjusted confidence interval limit is zero gives a rough adjusted *p*-value.

Sample size used

The sample size at the point the study was stopped should be reported in addition to the sample size that would have been used had the final stage been reached.

Technical Details

Many articles and texts have been written about group sequential analysis. Details of many of the relevant topics are discussed below, but this is not intended to be a comprehensive review of group-sequential methods. One of the more influential works in the area of group-sequential analysis is Jennison and Turnbull (2000).

Null and Alternative Hypotheses

For superiority by a margin tests of two means, the appropriate null and alternative hypotheses depend on whether higher means are better or higher means are worse.

Case 1: Low Values Good

In this case, higher values are worse. The hypotheses are arranged so that rejecting the null hypothesis implies that the treatment mean is less than the reference mean by at least the margin of superiority. The value of δ at which power is calculated must be less than $-|M_S|$. The null and alternative hypotheses with $\delta_0 = -|M_S|$ are

$H_0: \mu_1 \geq \mu_2 - |M_S|$ versus $H_1: \mu_1 < \mu_2 - |M_S|$

$H_0: \mu_1 - \mu_2 \geq -|M_S|$ versus $H_1: \mu_1 - \mu_2 < -|M_S|$

$H_0: \delta \geq -|M_S|$ versus $H_1: \delta < -|M_S|$

Case 2: High Values Good

In this case, higher values are better. The hypotheses are arranged so that rejecting the null hypothesis implies that the treatment mean is greater than the reference mean by at least the margin of superiority. The value of δ at which power is calculated must be greater than $|M_S|$. The null and alternative hypotheses with $\delta_0 = |M_S|$ are

$H_0: \mu_1 \leq \mu_2 + |M_S|$ versus $H_1: \mu_1 > \mu_2 + |M_S|$

$H_0: \mu_1 - \mu_2 \leq |M_S|$ versus $H_1: \mu_1 - \mu_2 > |M_S|$

$H_0: \delta \leq |M_S|$ versus $H_1: \delta > |M_S|$

Stages in Group-Sequential Testing

The potential to obtain the benefit from a group-sequential design and analysis occurs when the response data are collected over a period of weeks, months, or years rather than all at once. A typical example is the case where patients are enrolled in a study as they become available, as in many types of clinical trials.

A group-sequential testing stage is a point in the accumulation of the data where an interim analysis occurs, either by design or by necessity. At each stage, a test statistic is computed with all the accumulated data, and it is determined whether a boundary (efficacy or futility) is crossed. When an efficacy (or futility) boundary is crossed, the study is usually concluded, and inference is made. If the final stage is reached, the group-sequential design forces a decision of efficacy or futility at this stage.

For the discussions below, a non-specific interim analysis stage is referenced as k, and the final stage is K.

Test Statistic

The t-statistic for any stage k is obtained from all the accumulated data up to and including that stage.

t-Statistic Allowing Unequal Variance (Welch)

Welch's t-test is commonly used when no assumption of equal variances for the two populations is made. This is commonly known as the Aspin-Welch test, Welch's t-test (Welch, 1937), or the Satterthwaite method.

When lower means are better:

$$t_k = \frac{\bar{X}_{1k} - \bar{X}_{2k} - (-|SM|)}{\sqrt{\frac{s_{1k}^2}{n_{1k}} + \frac{s_{2k}^2}{n_{2k}}}} = \frac{\bar{X}_{1k} - \bar{X}_{2k} + |SM|}{\sqrt{\frac{s_{1k}^2}{n_{1k}} + \frac{s_{2k}^2}{n_{2k}}}}$$

When higher means are better:

$$t_k = \frac{\bar{X}_{1k} - \bar{X}_{2k} - |SM|}{\sqrt{\frac{s_{1k}^2}{n_{1k}} + \frac{s_{2k}^2}{n_{2k}}}}$$

with degrees of freedom

$$df = \frac{\left(\frac{s_{1k}^2}{n_{1k}} + \frac{s_{2k}^2}{n_{2k}}\right)^2}{\frac{\left(\frac{s_{1k}^2}{n_{1k}}\right)^2}{n_{1k} - 1} + \frac{\left(\frac{s_{2k}^2}{n_{2k}}\right)^2}{n_{2k} - 1}}$$

Group-Sequential Design Phase

In most group-sequential studies there is a design or planning phase prior to beginning response collection. In this phase, researchers specify the anticipated number and spacing of stages, the types of boundaries that will be used, the desired alpha and power levels, the spending functions, the anticipated standard deviations, and an estimate of the true difference in means.

Based on these input parameters, an initial set of boundaries is produced, an estimate of the total number of needed subjects is determined, and the anticipated total information at the final stage is calculated. The appropriate procedure in **PASS** (sample size software) can be used to make these planning phase sample size estimation calculations.

Information and Total Information

In the group-sequential design phase, the final stage (*K*) or total (design) information is calculated from the specified standard deviations and the final sample sizes, as

$$I_K^* = \frac{1}{\frac{\sigma_{1K}^2}{n_{1K}} + \frac{\sigma_{2K}^2}{n_{2K}}}$$

The (estimated) information at any stage *k* may be calculated from all the accumulated data up to and including that stage, as

$$I_k = \frac{1}{\frac{s_{1k}^2}{n_{1k}} + \frac{s_{2k}^2}{n_{2k}}}$$

The proportion of the total information (or information fraction) at any stage is

$$p_k = \frac{I_k}{I_K^*}$$

The information fractions are used in conjunction with the spending function(s) to define the alpha and/or beta to be spent at each stage.

To properly use the spending function at the final stage, it is required that $p_K = 1$. However, if the final stage is reached, we see that

$$I_K = \frac{1}{\frac{s_{1K}^2}{n_{1K\,achieved}} + \frac{s_{2K}^2}{n_{2K\,achieved}}} \neq I_K^* = \frac{1}{\frac{\sigma_{1K}^2}{n_{1K}} + \frac{\sigma_{2K}^2}{n_{2K}}}$$

so that

$$p_K = \frac{I_K}{I_K^*} \neq 1$$

When $I_K > I_K^*$, it is called over-running. When $I_K < I_K^*$, it is called under-running. In either case, the spending function is adjusted to accommodate the inequality, by redefining

$$I_K^* = I_K$$

See the discussion in Wassmer and Brannath (2016), pages 78-79, or Jennison and Turnbull (2000), pages 153-154, 162.

Types of Boundaries

A variety of boundary designs are available to reflect the needs of the study design.

Efficacy Only

The simplest group-sequential test involves a single set of stage boundaries with early stopping for efficacy.
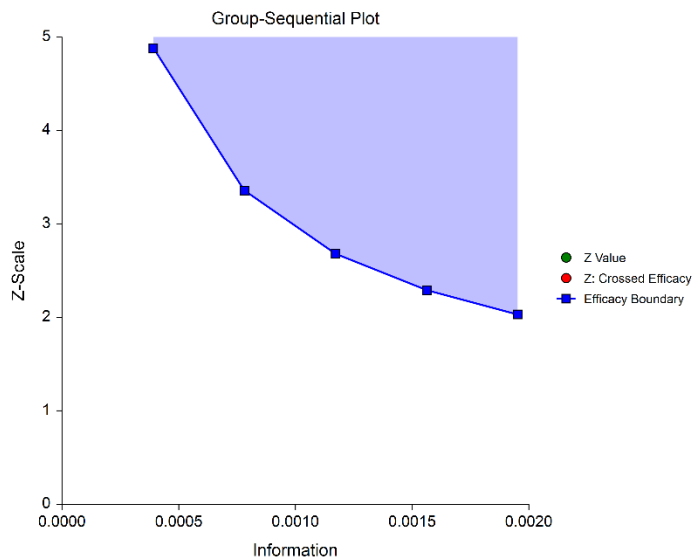
Efficacy and Binding Futility

This design allows early stopping for either efficacy or futility. For binding futility designs, the Type I error protection (alpha) is only maintained if the study is strictly required to stop if either boundary is crossed.
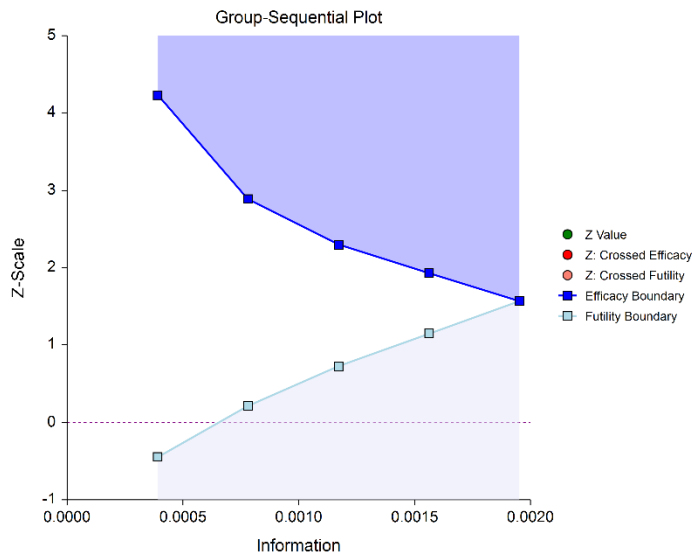
Efficacy and Non-Binding Futility

This design also allows early stopping for either efficacy or futility. For non-binding futility designs, the Type I error protection (alpha) is maintained, regardless of whether the study continues after crossing a futility boundary. However, the effect is to make the test conservative (alpha is lower than the stated alpha and power is lower than the stated power).
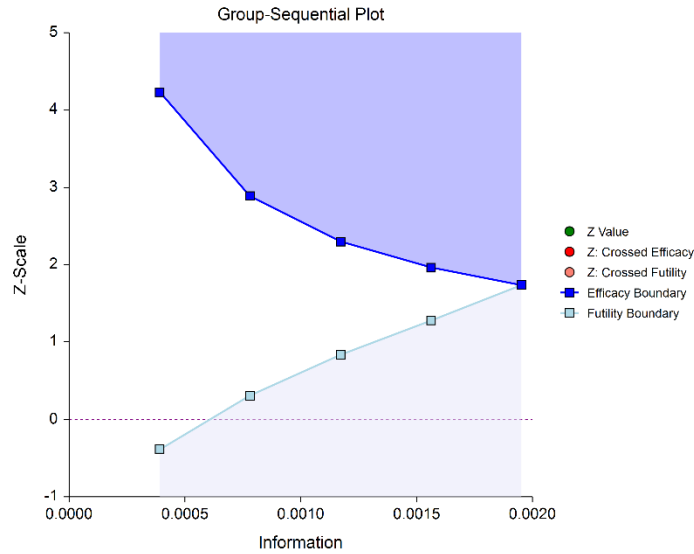
Futility Only (One-Sided)

In this design, the interim analyses are used only for futility. Please be aware that, due to computational complexity, these boundaries may take several minutes to compute, particularly when some stages are skipped.
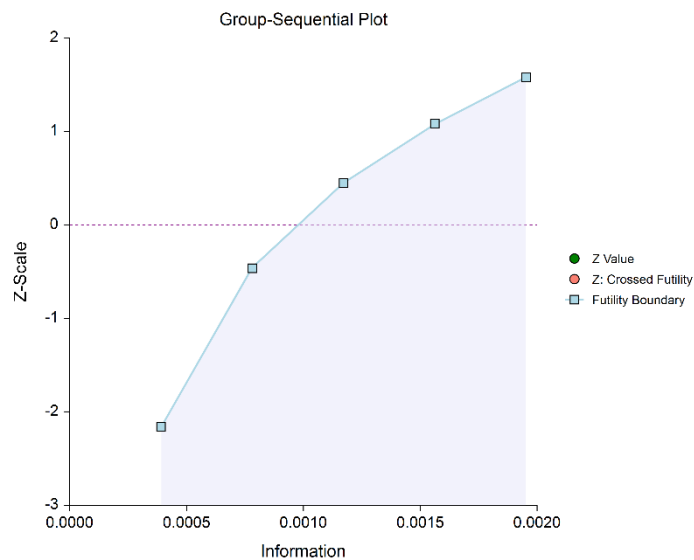
Boundary Calculations

The foundation of the spending function approach used in this procedure is given in Lan & DeMets (1983). This procedure implements the methods given in Reboussin, DeMets, Kim, & Lan (1992) to calculate the boundaries and stopping probabilities of the various group sequential designs. Some adjustments are made to these methods to facilitate the calculation of futility boundaries.

Obtaining t-scale Boundaries from Z-scale Boundaries

The methodology underlying all boundary calculations is based on *Z*-boundary calculations. Following the recommendation of Jennison and Turnbull (2000), p. 80, *Z*-scale boundaries are first produced, using the common methods. The corresponding *p*-value boundaries are computed from the *Z*-value boundaries. These *p*-value boundaries are then converted to *t*-value boundaries, using the appropriate degrees of freedom. As discussed in Jennison and Turnbull (2000), these *t*-value boundaries have been found to be 'remarkably accurate.'

Binding vs. Non-Binding Futility Boundaries

Futility boundaries are used to facilitate the early stopping of studies when early evidence leans to lack of efficacy. When binding futility boundaries are to be used, the calculation of the futility and efficacy boundaries assumes that the study will be strictly stopped at any stage where a futility or efficacy boundary is crossed. If strict adherence is not maintained, then the Type I and Type II error probabilities associated with the boundaries are no longer valid. One (perhaps undesirable) effect of using binding futility boundaries is that the resulting final stage boundary may be lower than the boundary given in the corresponding fixed-sample design.

When non-binding futility boundaries are calculated, the efficacy boundaries are first calculated ignoring futility boundaries completely. This is done so that alpha may be maintained whether or not a study continues after crossing a futility boundary. One (perhaps undesirable) effect of using non-binding futility boundaries is that the overall group-sequential test becomes conservative (alpha is lower than the stated alpha and power is lower than the stated power).

Spending Functions

Spending functions are used to distribute portions of alpha (or beta) to the stages according to the proportion of accumulated information at each look.

Spending Function Characteristics

- Spending functions give a value of zero when the proportion of accumulated information is zero.

$$\alpha(0) = 0 \text{ (for alpha-spending)}$$

$$\beta(0) = 0 \text{ (for beta-spending)}$$

- Spending functions are increasing functions.
- Spending functions give a value of alpha (or beta) when the proportion of accumulated information is one.

$$\alpha(1) = \alpha \text{ (for alpha-spending)}$$

$$\beta(1) = \beta \text{ (for beta-spending)}$$

Using spending functions in group-sequential analyses is very flexible in that neither the information proportions nor the number of stages need be specified in advance to maintain Type I and Type II error protection.

Spending Functions Available in this Procedure

The following spending functions are shown as alpha-spending functions. The corresponding beta-spending function is given by replacing α with β.

O'Brien-Fleming Analog

The O'Brien Fleming Analog (Lan & DeMets, 1983) roughly mimics the O'Brien-Fleming (non-spending function) design, with the key attribute that only a small proportion of alpha is spent early. Its popularity comes from it proportioning enough alpha to the final stage that the final stage boundary is not too different from the fixed-sample (non-group-sequential) boundary.

$$\alpha(0) = 0$$

$$\alpha(p_k) = 2 - 2\Phi\left(\frac{Z_{1-\alpha/2}}{\sqrt{p_k}}\right)$$

$$\alpha(1) = \alpha$$
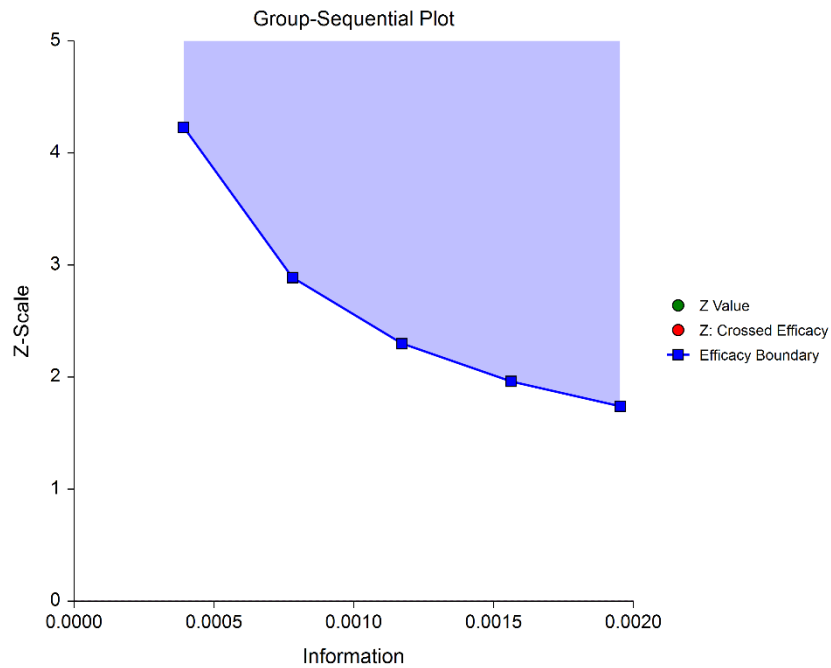
Pocock Analog

The Pocock Analog (Lan & DeMets, 1983) roughly mimics the Pocock (non-spending function) design, with the key attribute that alpha is spent roughly equally across all stages.

$$\alpha(0) = 0$$

$$\alpha(p_k) = \alpha\ln\left(1 + (e - 1)p_k\right)$$

$$\alpha(1) = \alpha$$
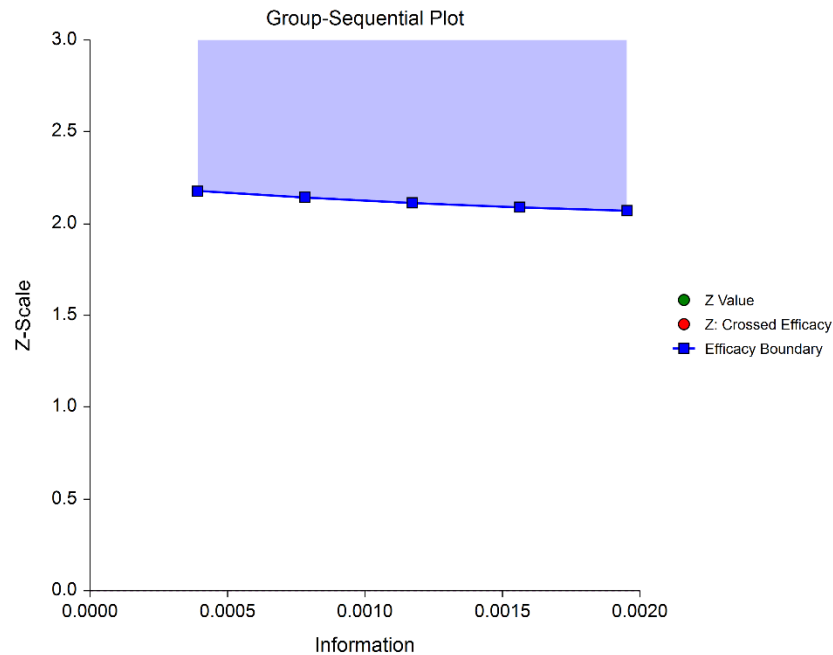
Power Family

The power family of spending functions has a ρ parameter that gives flexibility in the spending function shape.

$$\alpha(0) = 0$$

$$\alpha(p_k) = p_k^\rho, \ \ \rho > 0$$

$$\alpha(1) = \alpha$$

A power family spending function with a ρ of 1 is similar to a Pocock design, while a power family spending function with a ρ of 3 is more similar to an O'Brien-Fleming design.

Group-Sequential Superiority by a Margin T-Tests for Two Means

$\rho = 1$

$\rho = 2$
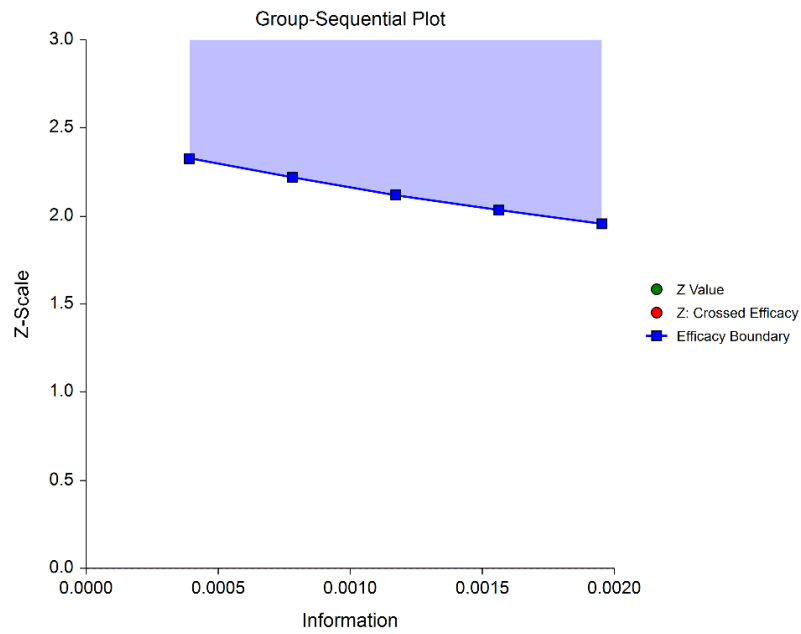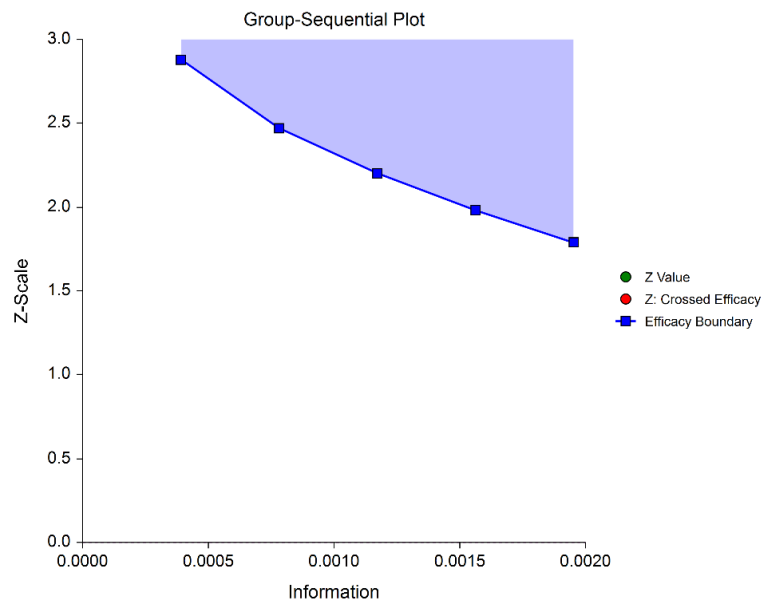
$\rho = 3$
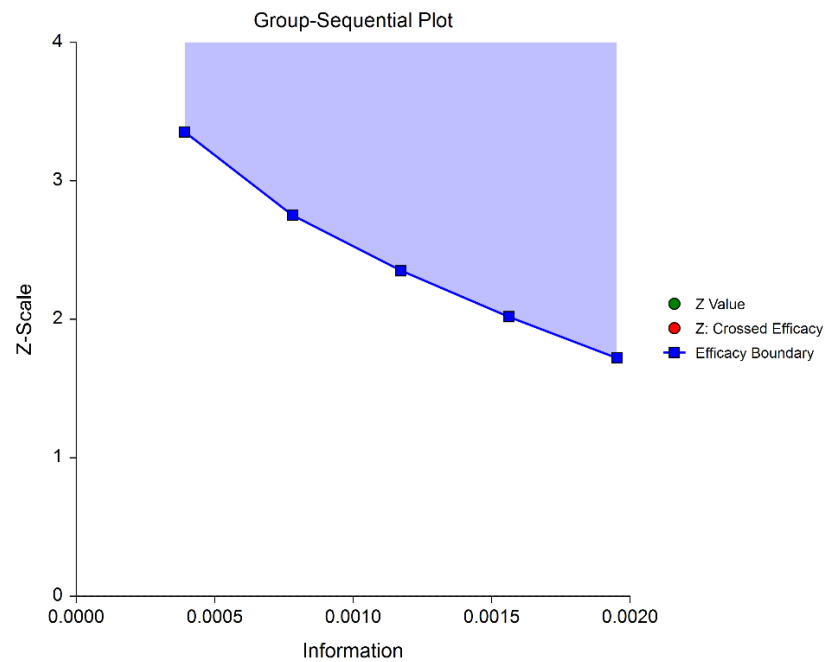
Hwang-Shih-DeCani (Gamma Family)

The Hwang-Shih-DeCani gamma family of spending function has a γ parameter that allows for a variety of spending functions.

$$\alpha(0) = 0$$

$$\alpha(p_k) = \alpha\left(\frac{1 - e^{-\gamma p_k}}{1 - e^{-\gamma}}\right), \ \gamma \neq 0$$

$$\alpha(p_k) = \alpha p_k, \ \gamma = 0$$

$$\alpha(1) = \alpha$$

Group-Sequential Superiority by a Margin T-Tests for Two Means

$\gamma = -3$

$\gamma = -1$
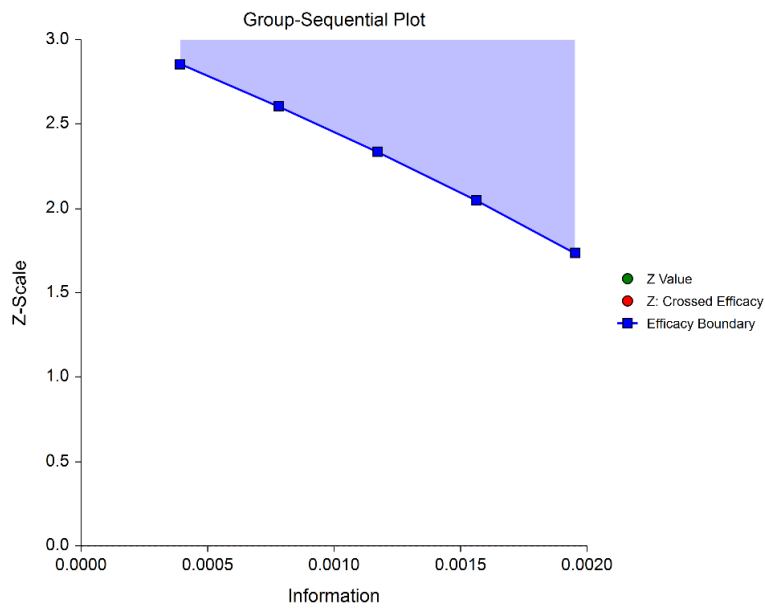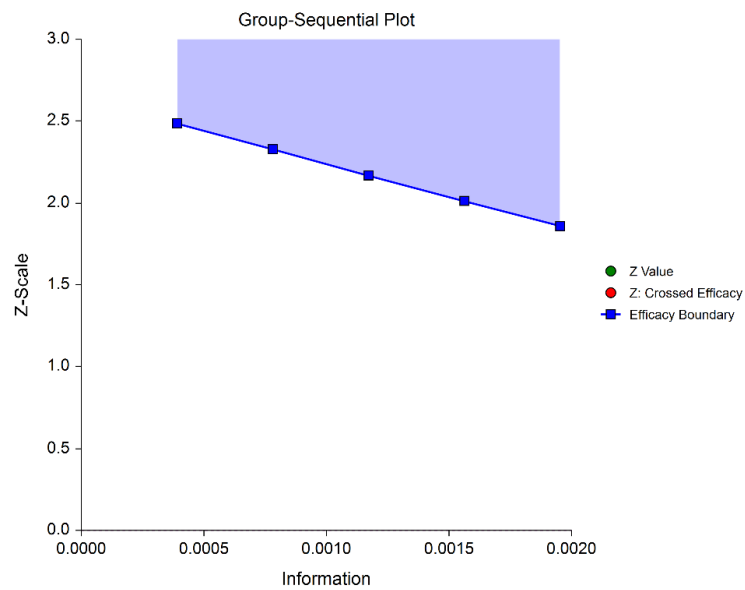
Group-Sequential Superiority by a Margin T-Tests for Two Means

$\gamma = 1$
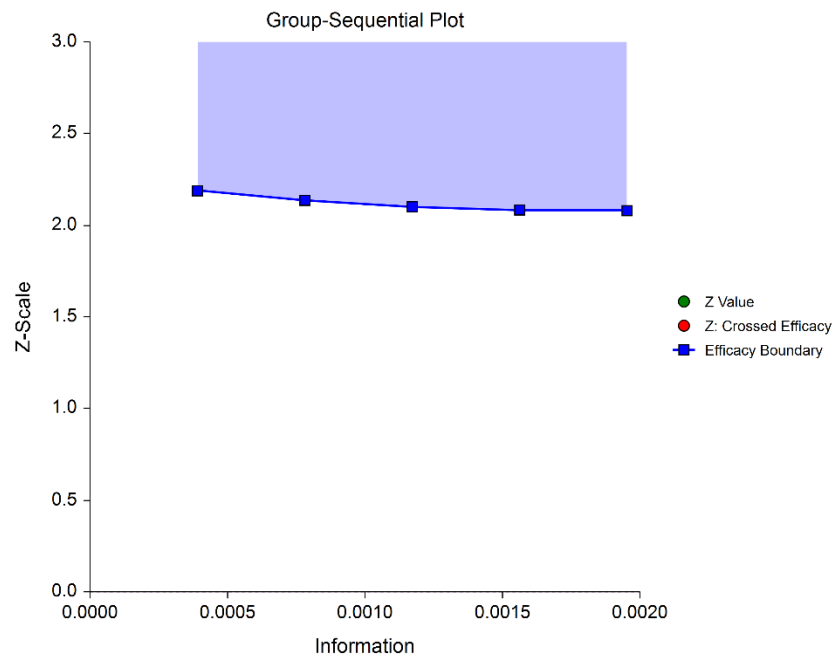
$\gamma = 3$
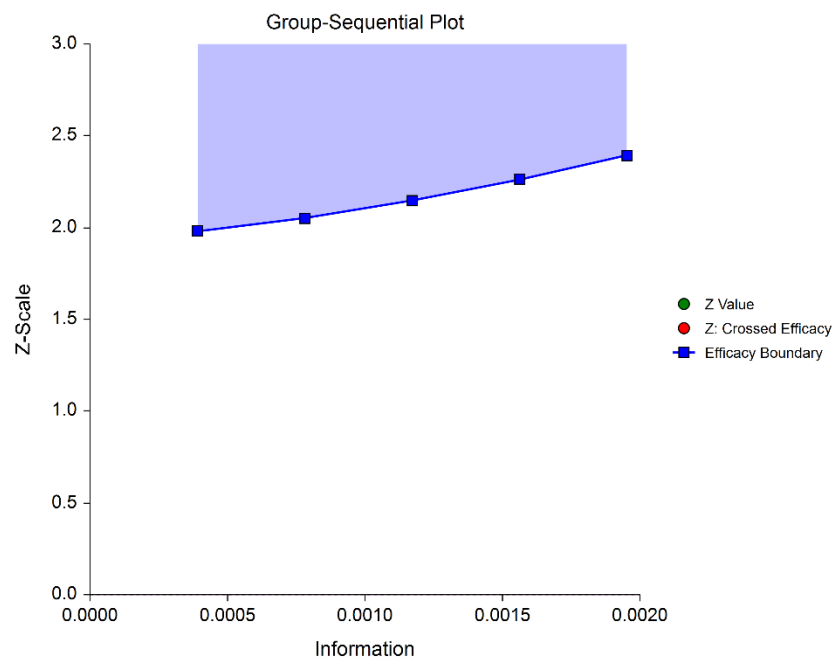
Adjusted Confidence Intervals

Except at the first stage, the raw (naïve) estimates of the confidence interval limits are inherently biased in the group-sequential analysis setting. The method given in Reboussin, DeMets, Kim, & Lan (1992) is used to calculate appropriately adjusted confidence limits and is based on Kim and DeMets (1987). This stage-wise ordering method is recommended by Jennison and Turnbull (2000) as "the only method available for use with unpredictable information sequences."

The methods used for the calculation of these confidence limits are based on the assumption that the current stage is the stopping stage of the study (typically from the crossing of a boundary).

As the methods in Reboussin *et al*. (1992) give only efficacy boundary results, the calculation adjustments are based only on the efficacy boundaries. The futility boundaries are not used in these calculations, except as they affect the efficacy boundaries.

Because these calculations are based on Z-scale methods, it is unclear to what degree the use of the *t*-statistic conversion affects the confidence interval accuracy.

Adjusted 'p-values'

As the raw (naïve) *p*-value is inherently biased in the group-sequential analysis setting, a sequential-test adjusted *p*-value should be used instead. Rather than giving a *p*-value in this procedure, a search is used to determine the adjusted confidence interval level at which an interval limit equals zero. As such, the methods used for this calculation are also based on the assumptions that the current stage is the stopping stage of the study. Again, only efficacy bounds are used.

Adjusted Mean Difference

The adjusted mean difference is a rough estimate of the difference calculated simply as the midpoint of the adjusted confidence interval limits.

Conditional Power

From Jennison and Turnbull (2000) pages 205 to 208, the general upper one-sided conditional power at stage *k* for rejecting a null hypothesis about a parameter θ at the end of the study, given the observed test statistic, Z_k, is computed as

$$P_{uk}(\theta) = \Phi\left(\frac{Z_k\sqrt{I_k} - z_{1-\alpha}\sqrt{I_K} + \theta(I_K - I_k)}{\sqrt{I_K - I_k}}\right),$$

the general lower one-sided conditional power at stage *k* is computed as

$$P_{lk}(\theta) = \Phi\left(\frac{-Z_k\sqrt{I_k} - z_{1-\alpha}\sqrt{I_K} - \theta(I_K - I_k)}{\sqrt{I_K - I_k}}\right),$$

and the general two-sided conditional power at stage k is computed as

$$P_k(\theta) = \Phi\left(\frac{Z_k\sqrt{I_k} - z_{1-\alpha/2}\sqrt{I_K} + \theta(I_K - I_k)}{\sqrt{I_K - I_k}}\right) + \Phi\left(\frac{-Z_k\sqrt{I_k} - z_{1-\alpha/2}\sqrt{I_K} - \theta(I_K - I_k)}{\sqrt{I_K - I_k}}\right),$$

where

  θ = the parameter being tested by the hypothesis

  k = an interim stage at which the conditional power is computed ($k = 1, …, K - 1$)

  K = the stage at which the study is terminated, and the final test computed

  Z_k = the test statistic calculated from the observed data that has been collected up to stage k

  I_k = the information level at stage k

  I_K = the information level at the end of the study

  $z_{1-\alpha}$ = the standard normal value for the test with a type I error rate of α.

For a test of a two means with null hypothesis $H_0: \mu_1 = \mu_2 \pm |SM|$, where μ_1 and μ_2 are the population means in groups 1 and 2, respectively, under the alternative hypothesis, these components are computed in Chang (2008) page 70 as

  $\theta = \mu_1 - \mu_2 \pm |SM|$ (the expected difference under the alternative hypothesis)

  $Z_k = (\bar{x}_{1k} - \bar{x}_{2k} \pm |SM|)\sqrt{\hat{I}_k}$ (the t statistic computed from the observed data)

  $I_k = \left(\frac{\sigma_1^2}{n_{1k}} + \frac{\sigma_2^2}{n_{2k}}\right)^{-1}$ (the interim information level)

  $I_K = \left(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)^{-1}$ (the final information level)

where

  $\bar{x}_{jk}$ is the sample mean for group j, estimating μ_j at stage k

  $\hat{I}_k$ is the estimated information from the sample at stage k

  n_{jk} is the sample size in group j at stage k

  n_j is the final sample size in group j

  σ_j^2 is the variance of group j

These methods for calculation of conditional power are based on Z-scale calculation with the substitution of sample standard deviations for known standard deviations. It is unclear to what degree the conversion to t-statistics affects the accuracy of the conditional power estimates.

Predictive Power

Predictive power (a Bayesian concept) is the result of averaging the conditional power over the posterior distribution of effect size. From Jennison and Turnbull (2000) pages 210 to 213, the general upper one-sided predictive power at stage *k* is given by

$$P_{uk} = \Phi\left(\frac{Z_k\sqrt{I_K} - z_{1-\alpha}\sqrt{I_k}}{\sqrt{I_K - I_k}}\right)$$

The general lower one-sided predictive power at stage *k* is given by

$$P_{lk} = \Phi\left(\frac{-Z_k\sqrt{I_K} - z_{1-\alpha}\sqrt{I_k}}{\sqrt{I_K - I_k}}\right)$$

The general two-sided predictive power at stage *k* is given by

$$P_k = \Phi\left(\frac{|Z_k|\sqrt{I_K} - z_{1-\alpha/2}\sqrt{I_k}}{\sqrt{I_K - I_k}}\right) + \Phi\left(\frac{-|Z_k|\sqrt{I_K} - z_{1-\alpha/2}\sqrt{I_k}}{\sqrt{I_K - I_k}}\right)$$

with all terms defined as in the equations for conditional power.

These methods for calculation of predictive power are based on *Z*-scale calculation with the substitution of sample standard deviations for known standard deviations. It is unclear to what degree the conversion to *t*-statistics affects the accuracy of the predictive power estimates.

Using Simulation to obtain Future Boundary Crossing Probabilities

It can be useful to researchers to know the probability of crossing future group-sequential boundaries, given the data already obtained, and given specified assumed values for the means. The following steps are used to estimate these probabilities using simulation:

1. From the data obtained to the current stage, obtain estimates of the standard deviations.

2. Determine the target (cumulative) sample sizes for each future stage, including the final stage. Fractional sample sizes are rounded up to the next integer.

3. For each simulation, append simulated values to the current sample data to obtain a data set with the final stage sample sizes. Simulated values correspond to the standard deviation estimates and the user-specified assumed means.

4. For each data set, determine which boundary or boundaries were crossed first (except in the case of non-binding futility boundaries). The proportion of simulations crossing each boundary (first) provides an estimate of the probability of crossing each boundary, given the specified assumed means.

Non-binding Futility Boundaries

When non-binding futility boundaries are used, the study may continue when a futility boundary is crossed. The simulation proportions will have a slightly different interpretation when this is the case.

Data Structure

The data for this procedure is entered in three columns. One column gives the individual responses. Another column identifies the group of the patient, individual, or experimental unit. A third column defines the associated stage of each response. The column for stages must use numeric integers beginning with 1 (1, 2, 3, ...).

The three columns are assigned on the Variables tab of the procedure. Groups 1 and 2 are also assigned on the Variables tab.

The software will assign the current stage as the highest value of the stage column. If the user wishes to exclude one or more (later) stages from the analysis, this may be done with the filter system.

Group-Sequential Analysis Data

Response	Group	Stage
44.2	Treatment	1
37.6	Placebo	1
65.8	Placebo	1
35.7	Treatment	1
53.4	Placebo	1
31.8	Treatment	1
38.4	Placebo	1
.	.	.
.	.	.
44.1	Placebo	2
49.5	Treatment	2
39.0	Treatment	2
54.7	Placebo	2
50.4	Placebo	2
61.2	Treatment	2
45.6	Placebo	2
36.8	Treatment	2
47.5	Treatment	2
.	.	.
.	.	.

Example 1 – Group-Sequential Analysis

A blood pressure treatment study is conducted to determine whether a new drug results in patient blood pressure readings that are lower than a current standard by at least 5. Thus, the desired margin of superiority is 5. The response for each patient is a resting systolic blood pressure. A one-sided test with alpha equal to 0.025 is used.

The new drug is assigned to Group 1, and the standard is assigned to Group 2, so that the null and alternative hypotheses are

$$H_0: \mu_1 - \mu_2 = -5 \quad (H_0: \mu_{New} = \mu_{Std} - 5)$$

versus

$$H_a: \mu_1 - \mu_2 < -5 \quad (H_a: \mu_{New} < \mu_{Std} - 5)$$

The design calls for five equally spaced stages if the final stage is reached. The current stage is the 3rd stage. In the design phase, a needed power of 0.90 called for 210 patients per group if the final stage is reached, based on known standard deviations of 22 for each group and assumed means of 112 and 124 for the new and standard treatments, respectively. Both efficacy and non-binding futility boundaries are implemented. The efficacy (alpha-spending) spending function used is the O'Brien-Fleming analog. The Hwang-Shih-DeCani (Gamma) beta-spending function with gamma parameter 1.5 is used for futility.

The boundary plot for stage 2 appeared as
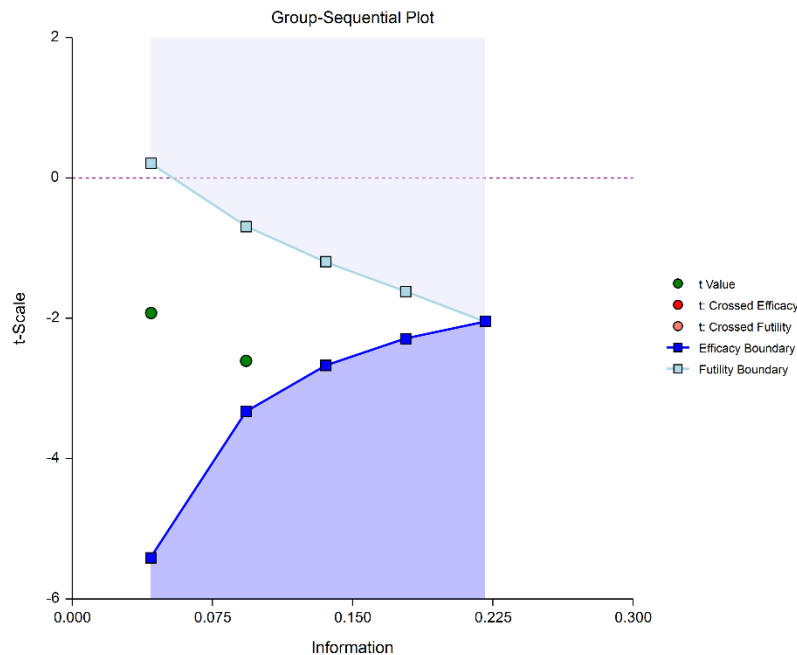
resulting in continuance of the study to stage 3.

Setup

To run this example, complete the following steps:

1   Open the BP SM example dataset

- From the File menu of the NCSS Data window, select **Open Example Data**.
- Select **BP SM** and click **OK**.

2   Specify the Group-Sequential Superiority by a Margin T-Tests for Two Means procedure options

- Find and open the **Group-Sequential Superiority by a Margin T-Tests for Two Means** procedure using the menus or the Procedure Navigator.
- The settings for this example are listed below and are stored in the **Example 1** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

Variables Tab

Response Variable	**Systolic_BP**
Group Variable..	**Treatment**
Group 1 Value..	**New**
Group 2 Value..	**Standard**
Stage Variable	**Stage**
Maximum Number of Stages (K)......................	**5**
Info. Proportion at each Stage	**Equally incremented**
N1	**210**
N2	**210**
µ1..	**112**
µ2..	**124**
σ10	**22**
σ20	**22**
Superiority Margin (SM)..................................	**5**
Future Stage Info. Adjustment	**Adjust information proportions proportional to original design**
Boundaries Used	**Efficacy with Futility**
Hypothesis Direction..	**Ha: µ1 - µ2 < -|SM| (Lower values are better)**
Boundary Specification	**Spending Function Calculation**
Alpha	**0.025**
Alpha Spending Function.................................	**O'Brien-Fleming Analog**
Skipped Efficacy Stages	**<Empty>**
Beta	**0.10**
Beta Spending Function	**Hwang-Shih-DeCani (γ)**
γ..	**1.5**
Skipped Futility Stages	**<Empty>**
Binding or Non-Binding Futility........................	**Non-Binding**

Reports Tab
_____

Reports for Current Stage

All Reports ....**Checked**

Confidence Level...**95**

All Three Boxes under Conditional ....**Checked**
and Predictive Power and Boundary
Crossing Probability Parameters

μ1...**120**

μ2...**126**

Planning Stage (Stage 0) Reports

Summary using t Scale...................................**Checked**

Plots..**Checked**

Plots Tab
_____

t-Statistic vs Information ....**Checked**

Note: You can add the stage number to the plot by checking the **small box on the Plot Format button** (under t-Statistic vs Information). When the procedure is run, select the **Stage Notes** tab and press the **Add Stage Number** button.

3 Run the procedure

- Click the **Run** button to perform the calculations and generate the output.

Run Summary Report

This report can be used to confirm that the input was processed as intended.

Run Summary Report
_____

Item	Value
Maximum Number of Stages (Design):	5
Current Stage:	3
Alternative Hypothesis:	μ1 - μ2 < -|SM|
Superiority Margin (SM):	5
t-Statistic:	t-Statistic Allowing Unequal Variance (Welch)
Alpha Spending Function:	O'Brien-Fleming Analog
Beta Spending Function:	Hwang-Shih-DeCani (γ = 1.5)
Futility Boundaries:	Non-Binding
Alpha:	0.0250
Maximum Information:	0.2169

_____

t-Values and Boundaries at Stage 3

This section gives the t-test values and boundaries, numerically. These values are reflected in the group-sequential boundary plot. The Decision column indicates whether a boundary was crossed at each stage.

t-Values and Boundaries at Stage 3

Maximum Information: 0.2169
Alternative Hypothesis: μ1 - μ2 < -|SM|
Superiority Margin (SM): 5
Futility Boundaries: Non-Binding

| | | Boundaries | | | |
| | t-Test | | | Information | |
Stage	Value	Efficacy	Futility	Proportion	Decision
1	-1.9252	-5.3532	0.1809	0.1946	Continue
2	-2.6069	-3.2919	-0.7303	0.4287	Continue
3	-3.1781	-2.2892	-1.7123	0.7906	Crossed Efficacy
4		-2.1966	-1.8030	0.8953	
5		-2.0806	-2.0806	1.0000	

Group-Sequential Boundary Plot at Stage 3

This plot shows the t-test values and boundaries. The efficacy boundary is crossed at Stage 3.

Group-Sequential Boundary Plot at Stage 3
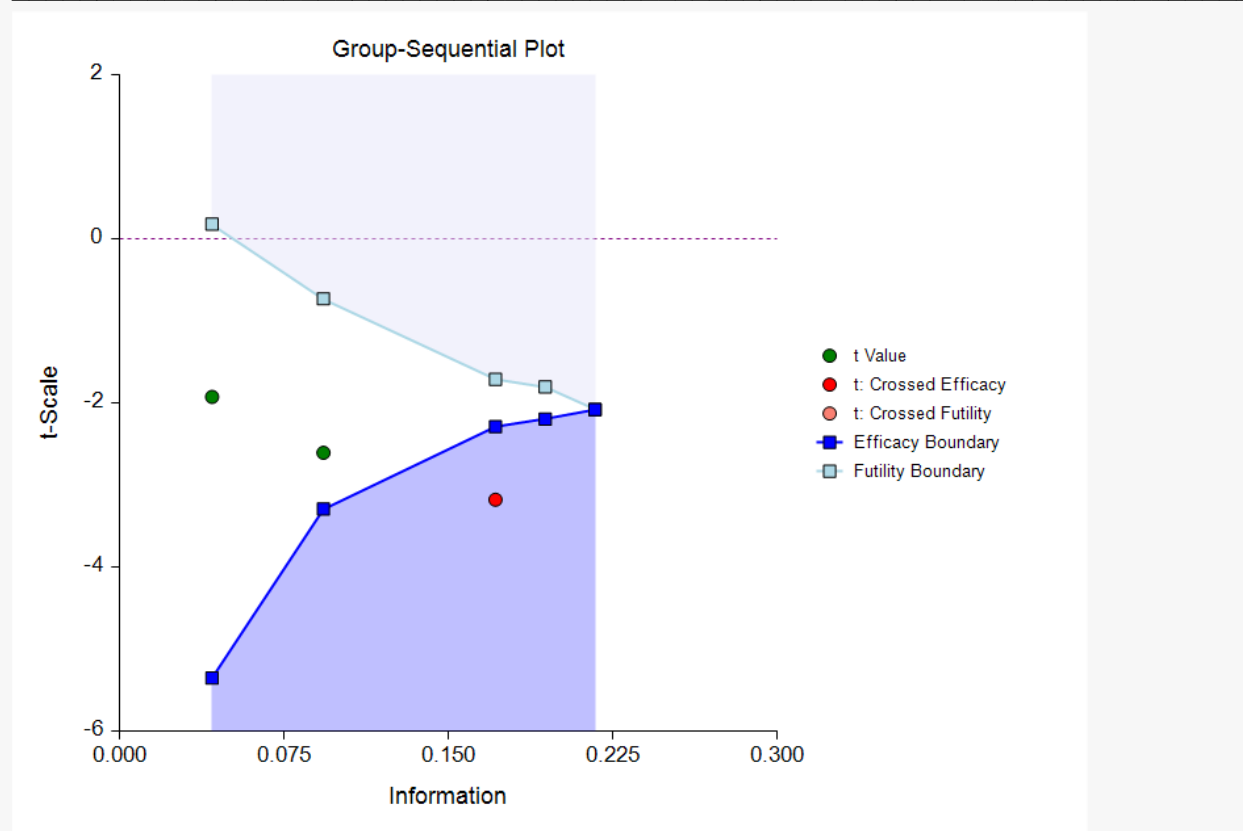
P-Values and Boundaries at Stage 3

This section reflects the conversion of the *Z*-test boundaries to the corresponding *P*-value boundaries. The *P*-value boundaries are then converted to the t-value boundaries used in the t-Values and Boundaries section. The T-Test P-Value is the raw *p*-value for the *t*-test at each stage.

P-Values and Boundaries at Stage 3

Maximum Information: 0.2169
Alternative Hypothesis: μ1 - μ2 < -|SM|
Superiority Margin (SM): 5
Futility Boundaries: Non-Binding

| | | Boundaries | | Information | |
| | t-Test | | | | |
Stage	P-Value	Efficacy	Futility	Proportion	Decision
1	0.02881	0.00000	0.57155	0.1946	Continue
2	0.00503	0.00062	0.23316	0.4287	Continue
3	0.00085	0.01150	0.04412	0.7906	Crossed Efficacy
4		0.01448	0.03629	0.8953	
5		0.01918	0.01918	1.0000	

P-values and P-value boundaries are one-sided values.

Underlying Z Scale Boundaries at Stage 3

This section gives the Z scale boundaries upon which the t scale boundaries are based. The transformation of Z scale boundaries to t scale boundaries is done by converting the Z boundaries to *p*-value boundaries and then converting the *p*-value boundaries to *t*-value boundaries, based on the corresponding degrees of freedom. Z-Test values are not given, since there are no available known standard deviations.

Underlying Z Scale Boundaries at Stage 3

Maximum Information: 0.2169
Alternative Hypothesis: μ1 - μ2 < -|SM|
Superiority Margin (SM): 5
Futility Boundaries: Non-Binding

| | | Boundaries | | Information | |
| | Z-Test | | | | |
Stage	Value	Efficacy	Futility	Proportion	Decision
1	N/A	-4.9483	0.1803	0.1946	see t-Test
2	N/A	-3.2300	-0.7285	0.4287	see t-Test
3	N/A	-2.2733	-1.7048	0.7906	see t-Test
4		-2.1841	-1.7955	0.8953	
5		-2.0709	-2.0709	1.0000	

Current Stage Results Adjusted for Sequential Analysis (Stage 3)

This section gives appropriate adjustments to the raw results to reflect the group-sequential nature of the analysis. Additional explanation is given in early sections of this chapter: Adjusted Confidence Intervals, Adjusted 'p-values', and Adjusted Mean Difference. Because these calculations are based on Z-scale methods, it is unclear to what degree the use of the *t*-statistic conversion affects the confidence interval accuracy.

Current Stage Results Adjusted for Sequential Analysis (Stage 3)

| | | Group-Sequential Adjusted Difference | | | Conf. Level Where Upper |
| | Actual | 95.0% C.I. of Diff. | | | Boundary |
Stage	Difference M1 - M2 + SM	Lower	Upper	C.I. Midpoint	Is Zero
3	-7.673905	-13.89198	-3.087912	-8.530404	99.752%

The adjustments are based only on the efficacy boundaries. The futility boundaries are not used in these calculations.

Descriptive Statistics up to Stage 3

This section gives the raw sample statistics for the cumulative data at each stage.

Descriptive Statistics up to Stage 3

| | Sample Size | | Mean | | Standard Deviation | | Difference | |
Stage	N1	N2	Mean 1	Mean 2	S1	S2	Difference	SE(Diff)
1	40	48	116.15	130.5208	18.57425	26.893	-14.37083	4.867487
2	82	85	112.8171	126.3647	17.37753	24.51183	-13.54763	3.27891
3	128	127	113.5859	126.2598	15.32972	22.52273	-12.67391	2.414584

Information Report at Stage 3

This section gives the target and achieved information for each stage, as well as the sample sizes and standard deviations used to calculate those informations. The degrees of freedom used at each stage is also reported.

Information Report at Stage 3

Maximum Information:	0.2169
Alternative Hypothesis:	μ1 - μ2 < -|SM|
Superiority Margin (SM):	5
Alpha:	0.0250

Stage	Information Proportion Target	Information Proportion Achieved	Information Target	Information Achieved	N1	N2	S1	S2	df
1	0.2000	0.1946	0.0434	0.0422	40.00	48.00	18.57425	26.893	83.31
2	0.4000	0.4287	0.0868	0.0930	82.00	85.00	17.37753	24.51183	151.64
3	0.6000	0.7906	0.1302	0.1715	128.00	127.00	15.32972	22.52273	221.93
4	0.8000	*0.8953	0.1736	*0.1942	*144.17	*144.17	*15.32972	*22.52273	*252.39
5	1.0000	*1.0000	0.2169	*0.2169	*161.03	*161.03	*15.32972	*22.52273	*282.10

* Projected value

Alpha Spending at Stage 3

This section shows how alpha was spent (or is anticipated to be spent) across the stages.

Alpha Spending at Stage 3

Target Final Stage Alpha:	0.0250
Spending Function:	O'Brien-Fleming Analog

Stage	Information Proportion	Alpha Spent this Stage	Cumulative Alpha Spent	Nominal (Boundary) Alpha	Percentage Alpha Spent this Stage	Cumulative Percentage Alpha Spent
1	0.1946	0.0000	0.0000	0.000000	0.0	0.0
2	0.4287	0.0006	0.0006	0.000619	2.5	2.5
3	0.7906	0.0111	0.0117	0.011503	44.4	46.8
4 *	0.8953	0.0061	0.0178	0.014478	24.5	71.4
5 *	1.0000	0.0072	0.0250	0.019185	28.6	100.0

* Projected

Beta Spending for Futility at Stage 3

This section shows how beta was spent (or is anticipated to be spent) across the stages.

Beta Spending for Futility at Stage 3

Target Cumulative Beta at Final Stage: 0.1000
Spending Function for Futility: Hwang-Shih-DeCani (γ = 1.5)

Stage	Information Proportion	Beta Spent this Stage	Cumulative Beta Spent	Nominal (Boundary) Beta	Percentage Beta Spent this Stage	Cumulative Percentage Beta Spent
1	0.1946	0.0326	0.0326	0.571552	32.6	32.6
2	0.4287	0.0285	0.0611	0.233157	28.5	61.1
3	0.7906	0.0283	0.0894	0.044120	28.3	89.4
4 *	0.8953	0.0057	0.0951	0.036287	5.7	95.1
5 *	1.0000	0.0049	0.1000	0.019185	4.9	100.0

* Projected

Conditional and Predictive Power Report at Stage 3

This section gives the conditional power for the various differences assumed. It also gives the predictive power.

Conditional and Predictive Power Report at Stage 3

Maximum Information: 0.2169
Information this Stage: 0.1715
Proportion of Maximum Information: 0.7906
Predictive Power: 0.9991

δ Name	δ Value	Conditional Power
Design	-12.0000	0.9996
Data	-12.6739	0.9998
δ1	-6.0000	0.9824

The conditional power values are calculated based on Jennison and Turnbull (2000), pages 205 to 208. They do not account for future interim stages, nor futility boundaries, if applicable. It is unclear to what degree the conversion to t-statistics affects the accuracy of the conditional and predictive power estimates.

Boundary Probabilities for δ = -12

Using simulation based on the specified means and standard deviations, this section gives the estimated probabilities of crossing each of the future boundaries. Values given here will vary for each simulation.

Boundary Probabilities for δ = -12
───

Number of Simulations:	2000
User-Entered Random Seed:	7883448
Futility Boundaries:	Non-Binding
After Efficacy Boundary Crossing:	Hold Out
After Non-Binding Futility Boundary Crossing:	Leave In
Alternative Hypothesis:	μ1 - μ2 < -|SM|
Superiority Margin (SM):	5
μ1:	112
μ2:	124
δ:	-12
σ1:	15.3297237617067
σ2:	22.5227266521622

				Efficacy		Futility	
Stage	N1	N2	t-Test Value	Boundary	Probability	Boundary	Probability
1	40.00	48.00	-1.9252	-5.3532		0.1809	
2	82.00	85.00	-2.6069	-3.2919		-0.7303	
3	128.00	127.00	-3.1781	-2.2892		-1.7123	
4	*144.17	*144.17		-2.1966	1.0000	-1.8030	0.0000
5	*161.03	*161.03		-2.0806	0.0000	-2.0806	0.0005

* Simulation sample size (Non-integer sample sizes were rounded to the next highest integer.)

Boundary Probabilities for δ = -12.67391

This section estimates the probabilities of crossing future boundaries if the current sample means are assumed. Values given here will vary for each simulation.

Boundary Probabilities for δ = -12.67391
_____

Number of Simulations:	2000
User-Entered Random Seed:	7883448
Futility Boundaries:	Non-Binding
After Efficacy Boundary Crossing:	Hold Out
After Non-Binding Futility Boundary Crossing:	Leave In
Alternative Hypothesis:	μ1 - μ2 < -|SM|
Superiority Margin (SM):	5
μ1:	113.5859
μ2:	126.2598
δ:	-12.67391
σ1:	15.3297237617067
σ2:	22.5227266521622

Stage	N1	N2	t-Test Value	Efficacy Boundary	Efficacy Probability	Futility Boundary	Futility Probability
1	40.00	48.00	-1.9252	-5.3532		0.1809	
2	82.00	85.00	-2.6069	-3.2919		-0.7303	
3	128.00	127.00	-3.1781	-2.2892		-1.7123	
4	*144.17	*144.17		-2.1966	1.0000	-1.8030	0.0000
5	*161.03	*161.03		-2.0806	0.0000	-2.0806	0.0015

* Simulation sample size (Non-integer sample sizes were rounded to the next highest integer.)

Boundary Probabilities for δ = -6

This section estimates the probabilities of crossing future boundaries if the difference for the remaining looks is assumed to be 2.

Boundary Probabilities for δ = -6
───

Number of Simulations:	2000
User-Entered Random Seed:	7883448
Futility Boundaries:	Non-Binding
After Efficacy Boundary Crossing:	Hold Out
After Non-Binding Futility Boundary Crossing:	Leave In
Alternative Hypothesis:	μ1 - μ2 < -|SM|
Superiority Margin (SM):	5
μ1:	120
μ2:	126
δ:	-6
σ1:	15.3297237617067
σ2:	22.5227266521622

Stage	N1	N2	t-Test Value	Efficacy Boundary	Efficacy Probability	Futility Boundary	Futility Probability
1	40.00	48.00	-1.9252	-5.3532		0.1809	
2	82.00	85.00	-2.6069	-3.2919		-0.7303	
3	128.00	127.00	-3.1781	-2.2892		-1.7123	
4	*144.17	*144.17		-2.1966	0.9895	-1.8030	0.0005
5	*161.03	*161.03		-2.0806	0.0050	-2.0806	0.0345

* Simulation sample size (Non-integer sample sizes were rounded to the next highest integer.)

t-Values and Boundaries at Stage 0

This section gives the boundaries that were initially projected at the planning stage.

t-Values and Boundaries at Stage 0
───

Maximum Information:	0.2169
Alternative Hypothesis:	μ1 - μ2 < -|SM|
Superiority Margin (SM):	5
Futility Boundaries:	Non-Binding

Stage	t-Test Value	Boundaries Efficacy	Boundaries Futility	Information Proportion	Decision
1		-5.2709	0.1539	0.2000	
2		-3.4200	-0.5995	0.4000	
3		-2.7024	-1.1569	0.6000	
4		-2.3005	-1.6054	0.8000	
5		-2.0372	-2.0372	1.0000	

Group-Sequential Boundary Plot at Planning Stage (Stage 0)

This plot shows the projected boundaries at the planning stage.

Group-Sequential Boundary Plot at Planning Stage (Stage 0)
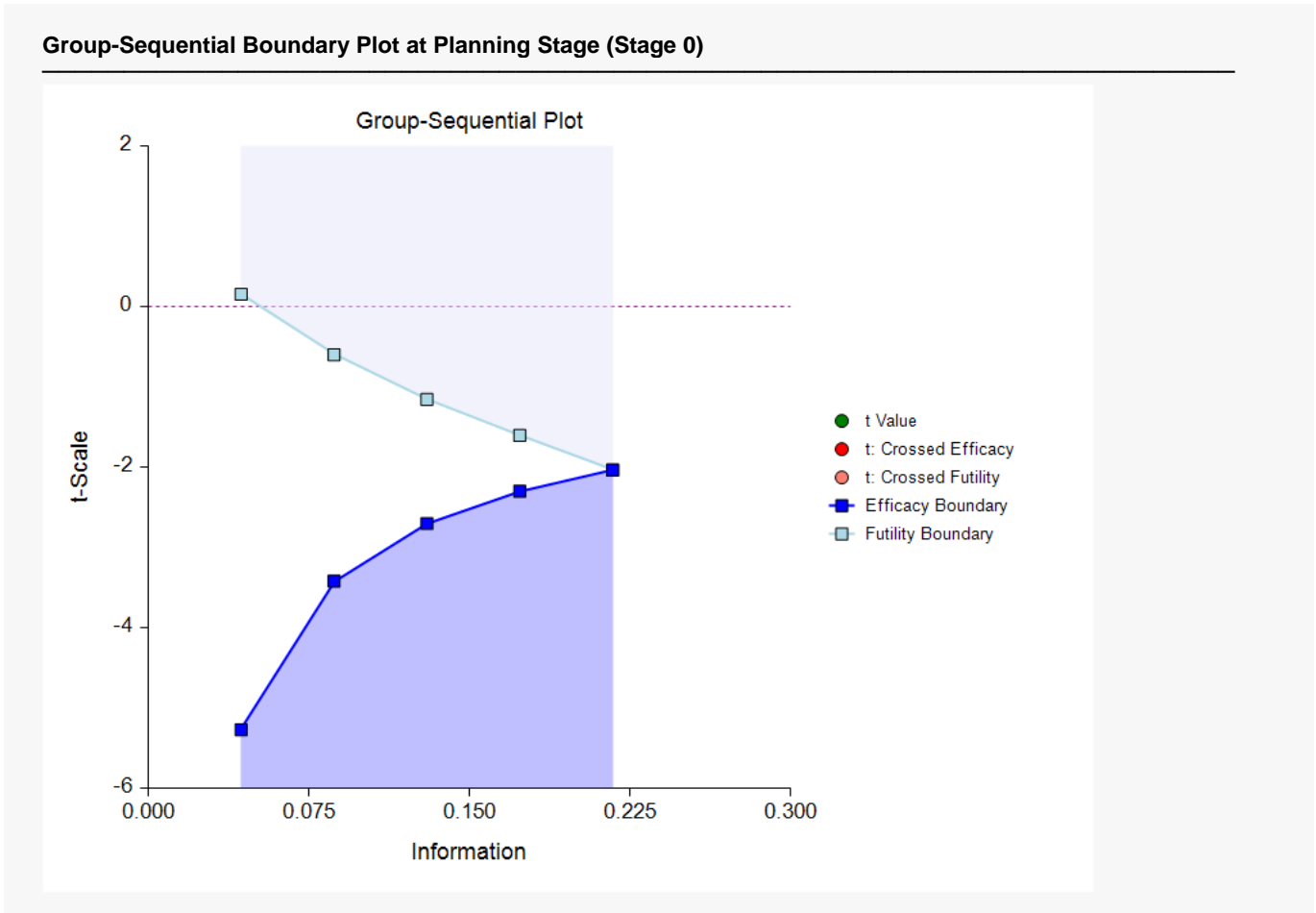
Example 2 – Skipping Stage Boundaries

Suppose that the setup is the same as in Example 1, except that the first two futility boundaries are skipped.

Setup

To run this example, complete the following steps:

1   Open the BP SM example dataset
- From the File menu of the NCSS Data window, select **Open Example Data**.
- Select **BP SM** and click **OK**.

2   Specify the Group-Sequential Superiority by a Margin T-Tests for Two Means procedure options
- Find and open the **Group-Sequential Superiority by a Margin T-Tests for Two Means** procedure using the menus or the Procedure Navigator.
- The settings for this example are listed below and are stored in the **Example 2** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

  Variables Tab

  Skipped Futility Stages **1 2**

3   Run the procedure
- Click the **Run** button to perform the calculations and generate the output.

Run Summary Report

The skipped futility stages are now reported.

Run Summary Report

Item	Value
Maximum Number of Stages (Design):	5
Skipped Futility Stage(s):	1 2
Current Stage:	3
Alternative Hypothesis:	$\mu_1 - \mu_2 < -|SM|$
Superiority Margin (SM):	5
t-Statistic:	t-Statistic Allowing Unequal Variance (Welch)
Alpha Spending Function:	O'Brien-Fleming Analog
Beta Spending Function:	Hwang-Shih-DeCani ($\gamma = 1.5$)
Futility Boundaries:	Non-Binding
Alpha:	0.0250
Maximum Information:	0.2169

t-Values and Boundaries at Stage 3

The boundaries change slightly from those where no boundaries are skipped.

**t-Values and Boundaries at Stage 3**

Maximum Information: 0.2169
Alternative Hypothesis: $\mu_1 - \mu_2 < -|SM|$
Superiority Margin (SM):   5
Futility Boundaries: Non-Binding

Stage	t-Test Value	Boundaries Efficacy	Futility	Information Proportion	Decision
1	-1.9252	-5.3532		0.1946	Continue
2	-2.6069	-3.2919		0.4287	Continue
3	-3.1781	-2.2892	-1.8818	0.7906	Crossed Efficacy
4		-2.1966	-1.8373	0.8953	
5		-2.0806	-2.0806	1.0000	

Group-Sequential Boundary Plot at Stage 3

The plot now has two futility boundaries skipped.

Group-Sequential Boundary Plot at Stage 3
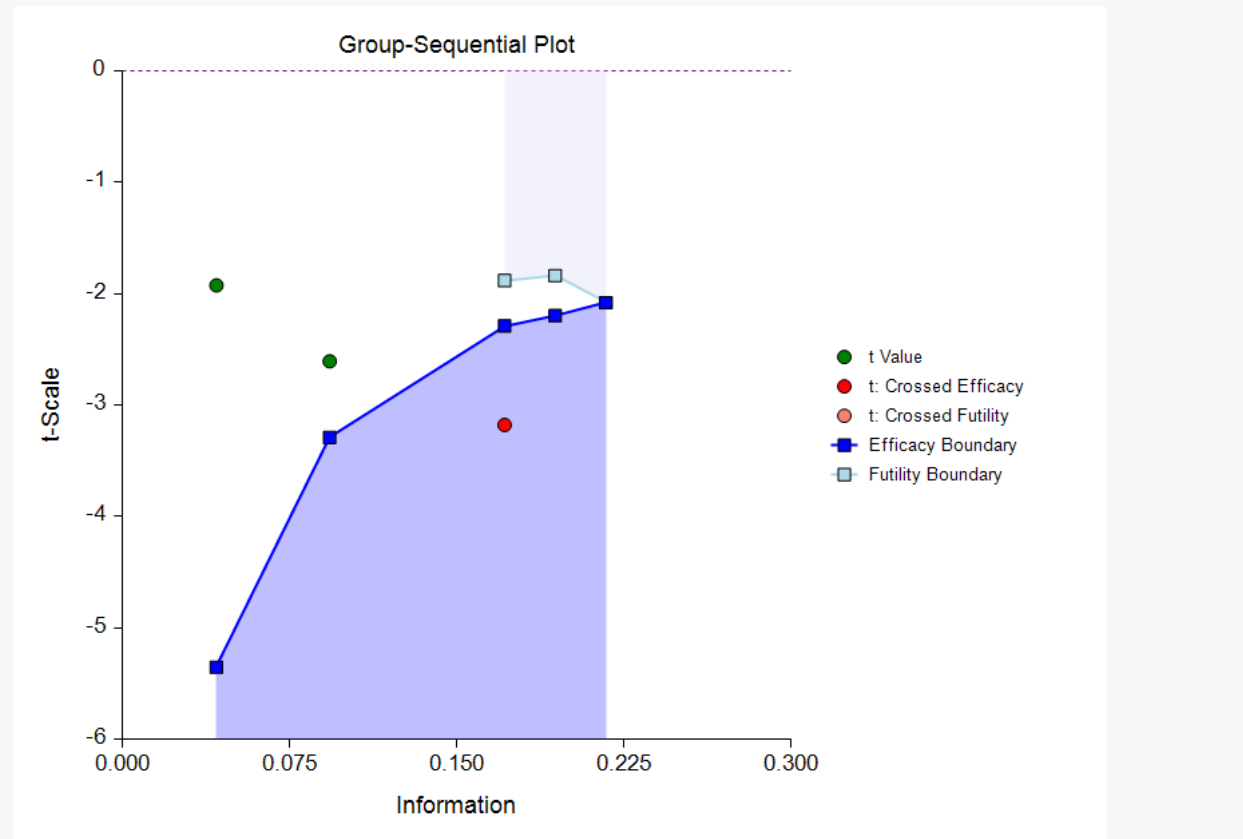
Example 3 – Sample Size Re-estimation and Boundary Probabilities

Suppose that the setup is the same as in Example 1, except that only the data up through Stage 2 has been collected.

You may follow along here by making the appropriate entries or load the completed template **Example 3** by clicking on Open Example Template from the File menu on the procedure window.

Setup

To run this example, complete the following steps:

1 Open the BP SM 2 Stages example dataset
- From the File menu of the NCSS Data window, select **Open Example Data**.
- Select **BP SM 2 Stages** and click **OK**.

2 Specify the Group-Sequential Superiority by a Margin T-Tests for Two Means procedure options
- Find and open the **Group-Sequential Superiority by a Margin T-Tests for Two Means** procedure using the menus or the Procedure Navigator.
- The settings for this example are stored in the **Example 3** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

3 Run the procedure
- Click the **Run** button to perform the calculations and generate the output.

Run Summary Report

This report can be used to confirm that the input was processed as intended.

Run Summary Report

Item	Value
Maximum Number of Stages (Design):	5
Current Stage:	2
Alternative Hypothesis:	$\mu 1 - \mu 2 < -|SM|$
Superiority Margin (SM):	5
t-Statistic:	t-Statistic Allowing Unequal Variance (Welch)
Alpha Spending Function:	O'Brien-Fleming Analog
Beta Spending Function:	Hwang-Shih-DeCani ($\gamma = 1.5$)
Futility Boundaries:	Non-Binding
Alpha:	0.0250
Maximum Information:	0.2169

t-Values and Boundaries at Stage 2

Examining the t-Tests values, the boundaries, and the decisions, no boundary has been crossed at this stage.

t-Values and Boundaries at Stage 2

Maximum Information: 0.2169
Alternative Hypothesis: μ1 - μ2 < -|SM|
Superiority Margin (SM): 5
Futility Boundaries: Non-Binding

| | | Boundaries | | | |
| | t-Test | --------------------- | | Information | |
Stage	Value	Efficacy	Futility	Proportion	Decision
1	-1.9252	-5.3532	0.1854	0.1946	Continue
2	-2.6069	-3.2919	-0.7236	0.4287	Continue
3		-2.6606	-1.2064	0.6192	
4		-2.2908	-1.6256	0.8096	
5		-2.0422	-2.0422	1.0000	

Group-Sequential Boundary Plot at Stage 2

This plot shows the t-test values approaching the efficacy boundary side.

Group-Sequential Boundary Plot at Stage 2

P-Values and Boundaries at Stage 2

This section reflects the conversion of the *Z*-test boundaries to the corresponding *P*-value boundaries. The *P*-value boundaries are then converted to the t-value boundaries used in the t-Values and Boundaries section.

P-Values and Boundaries at Stage 2
───

Maximum Information: 0.2169
Alternative Hypothesis: $\mu_1 - \mu_2 < -|SM|$
Superiority Margin (SM): 5
Futility Boundaries: Non-Binding
───

| | | **Boundaries** | | | |
| | t-Test | ─────────────── | | Information | |
Stage	P-Value	Efficacy	Futility	Proportion	Decision
1	0.02881	0.00000	0.57333	0.1946	Continue
2	0.00503	0.00062	0.23522	0.4287	Continue
3		0.00419	0.11449	0.6192	
4		0.01136	0.05258	0.8096	
5		0.02094	0.02094	1.0000	

P-values and P-value boundaries are one-sided values.

Underlying Z Scale Boundaries at Stage 2

This section gives the Z scale boundaries upon which the t scale boundaries are based. The transformation of Z scale boundaries to t scale boundaries is done by converting the Z boundaries to *p*-value boundaries and then converting the *p*-value boundaries to *t*-value boundaries, based on the corresponding degrees of freedom. Z-Test values are not given, since there are no available known standard deviations.

Underlying Z Scale Boundaries at Stage 2
───
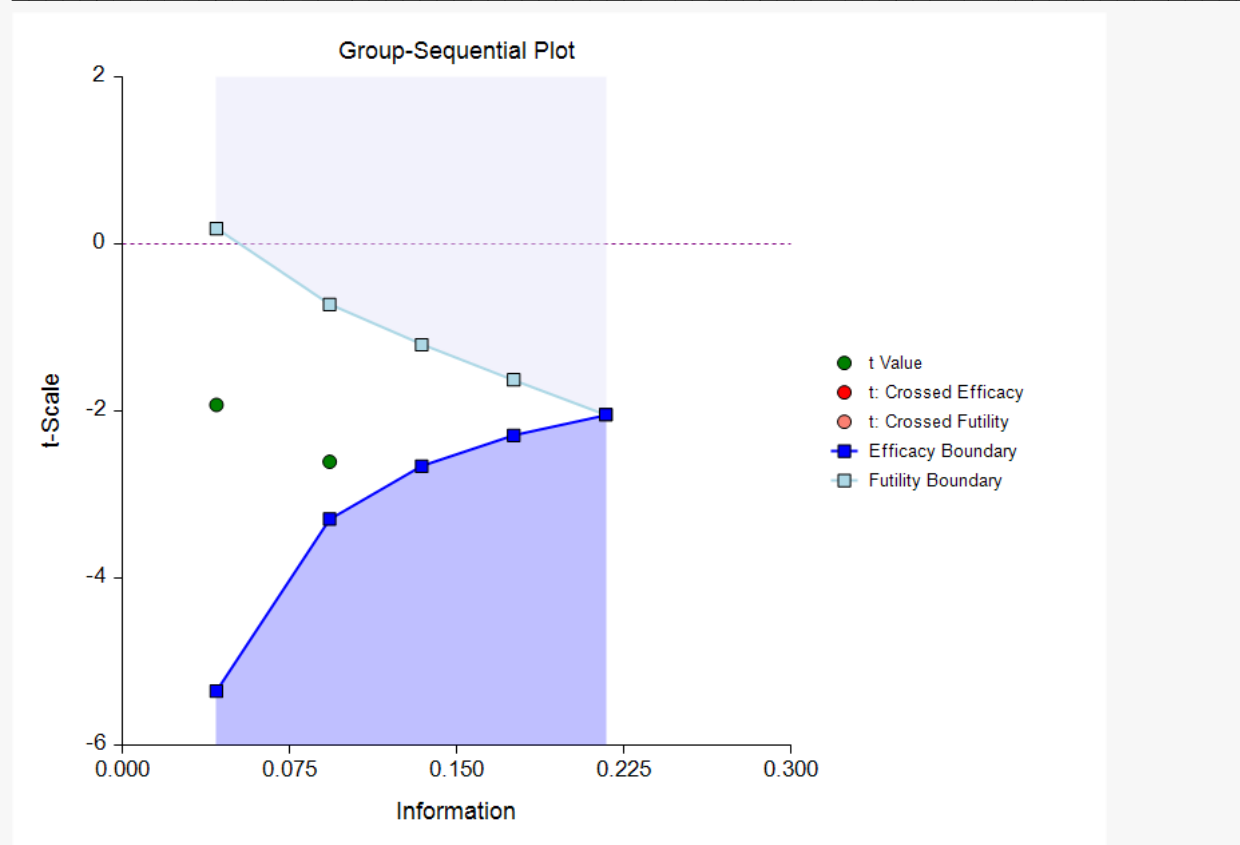
Maximum Information: 0.2169
Alternative Hypothesis: $\mu_1 - \mu_2 < -|SM|$
Superiority Margin (SM): 5
Futility Boundaries: Non-Binding
───

| | | **Boundaries** | | | |
| | Z-Test | ─────────────── | | Information | |
Stage	Value	Efficacy	Futility	Proportion	Decision
1	N/A	-4.9483	0.1849	0.1946	see t-Test
2	N/A	-3.2300	-0.7218	0.4287	see t-Test
3		-2.6362	-1.2030	0.6192	
4		-2.2783	-1.6204	0.8096	
5		-2.0347	-2.0347	1.0000	

Current Stage Results Adjusted for Sequential Analysis (Stage 2)

This section gives appropriate adjustments to the raw results to reflect the group-sequential nature of the analysis.

Current Stage Results Adjusted for Sequential Analysis (Stage 2)

| | | Group-Sequential Adjusted Difference | | | Conf. Level Where Upper Boundary Is Zero |
| | Actual Difference M1 - M2 + SM | 95.0% C.I. of Diff. | | C.I. Midpoint | |
Stage		Lower	Upper		
2	-8.547633	-22.86885	-3.239366	-13.05411	99.086%

Descriptive Statistics up to Stage 2

This section gives the raw sample statistics for the cumulative data at each stage.

Descriptive Statistics up to Stage 2

| | Sample Size | | Mean | | Standard Deviation | | Difference | |
Stage	N1	N2	Mean 1	Mean 2	S1	S2	Difference	SE(Diff)
1	40	48	116.15	130.5208	18.57425	26.893	-14.37083	4.867487
2	82	85	112.8171	126.3647	17.37753	24.51183	-13.54763	3.27891

Information Report at Stage 2 (Gives Sample Size Re-estimation)

This section shows that the target sample size for the next stage should be 122 per group, and the final stage sample size is now estimated at 196 per group. The final stage sample size can be compared to the original estimate of 213 per group, where a standard deviation of 22 for each group was assumed.

Information Report at Stage 2

Maximum Information: 0.2169
Alternative Hypothesis: $\mu_1 - \mu_2 < -|SM|$
Superiority Margin (SM): 5
Alpha: 0.0250

| | Information Proportion | | Information | | | | | | |
Stage	Target	Achieved	Target	Achieved	N1	N2	S1	S2	df
1	0.2000	0.1946	0.0434	0.0422	40.00	48.00	18.57425	26.893	83.31
2	0.4000	0.4287	0.0868	0.0930	82.00	85.00	17.37753	24.51183	151.64
3	0.6000	*0.6192	0.1302	*0.1343	*121.27	*121.27	*17.37753	*24.51183	*216.78
4	0.8000	*0.8096	0.1736	*0.1756	*158.56	*158.56	*17.37753	*24.51183	*284.00
5	1.0000	*1.0000	0.2169	*0.2169	*195.86	*195.86	*17.37753	*24.51183	*351.23

* Projected value

Alpha Spending at Stage 2

This section shows how alpha was spent (or is anticipated to be spent) across the stages.

Alpha Spending at Stage 2

Target Final Stage Alpha: 0.0250
Spending Function: O'Brien-Fleming Analog

Stage	Information Proportion	Alpha Spent this Stage	Cumulative Alpha Spent	Nominal (Boundary) Alpha	Percentage Alpha Spent this Stage	Cumulative Percentage Alpha Spent
1	0.1946	0.0000	0.0000	0.000000	0.0	0.0
2	0.4287	0.0006	0.0006	0.000619	2.5	2.5
3 *	0.6192	0.0038	0.0044	0.004192	15.1	17.6
4 *	0.8096	0.0083	0.0127	0.011355	33.4	50.9
5 *	1.0000	0.0123	0.0250	0.020940	49.1	100.0

* Projected

Beta Spending for Futility at Stage 2

This section shows how beta was spent (or is anticipated to be spent) across the stages.

Beta Spending for Futility at Stage 2

Target Cumulative Beta at Final Stage: 0.1000
Spending Function for Futility: Hwang-Shih-DeCani (γ = 1.5)

Stage	Information Proportion	Beta Spent this Stage	Cumulative Beta Spent	Nominal (Boundary) Beta	Percentage Beta Spent this Stage	Cumulative Percentage Beta Spent
1	0.1946	0.0326	0.0326	0.573331	32.6	32.6
2	0.4287	0.0285	0.0611	0.235222	28.5	61.1
3 *	0.6192	0.0168	0.0779	0.114489	16.8	77.9
4 *	0.8096	0.0126	0.0905	0.052575	12.6	90.5
5 *	1.0000	0.0095	0.1000	0.020940	9.5	100.0

* Projected

Conditional and Predictive Power Report at Stage 2

Conditional power and predictive power are described earlier in the Technical details section. The predictive power does not depend on an assumed difference, as does the conditional power.

Conditional and Predictive Power Report at Stage 2
───

Maximum Information:	0.2169
Information this Stage:	0.0930
Proportion of Maximum Information:	0.4287
Predictive Power:	0.9600

───

δ Name	δ Value	Conditional Power
Design	-12.0000	0.9834
Data	-13.5476	0.9963
δ1	-6.0000	0.5069

───

The conditional power values are calculated based on Jennison and Turnbull (2000), pages 205 to 208. They do not account for future interim stages, nor futility boundaries, if applicable. It is unclear to what degree the conversion to t-statistics affects the accuracy of the conditional and predictive power estimates.

Boundary Probabilities for δ = -12

Given the data that have already accumulated in the first two stages, and assuming means of 112 and 124 and standard deviations of 17.376 and 24.512 going forward, this report gives the simulation probabilities of crossing each of the future boundaries. The sum of the efficacy boundary probabilities is another estimate of the conditional power.

Boundary Probabilities for δ = -12
───

Number of Simulations:	2000
User-Entered Random Seed:	7970145
Futility Boundaries:	Non-Binding
After Efficacy Boundary Crossing:	Hold Out
After Non-Binding Futility Boundary Crossing:	Leave In
Alternative Hypothesis:	μ1 - μ2 < -|SM|
Superiority Margin (SM):	5
μ1:	112
μ2:	124
δ:	-12
σ1:	17.3775277395911
σ2:	24.5118275915271

───

			t-Test	Efficacy		Futility	
Stage	N1	N2	Value	Boundary	Probability	Boundary	Probability
1	40.00	48.00	-1.9252	-5.3532		0.1854	
2	82.00	85.00	-2.6069	-3.2919		-0.7236	
3	*121.27	*121.27		-2.6606	0.7135	-1.2064	0.0010
4	*158.56	*158.56		-2.2908	0.2220	-1.6256	0.0070
5	*195.86	*195.86		-2.0422	0.0485	-2.0422	0.0220

───

* Simulation sample size (Non-integer sample sizes were rounded to the next highest integer.)

Boundary Probabilities for δ = -13.54763

Given the data that have already accumulated in the first two stages, and assuming means of 112.8171 and 1263647 and standard deviations of 17.376 and 24.512 going forward, this report gives the simulation probabilities of crossing each of the future boundaries. The sum of the efficacy boundary probabilities is another estimate of the conditional power.

Boundary Probabilities for δ = -13.54763
―――

Number of Simulations:	2000
User-Entered Random Seed:	7970145
Futility Boundaries:	Non-Binding
After Efficacy Boundary Crossing:	Hold Out
After Non-Binding Futility Boundary Crossing:	Leave In
Alternative Hypothesis:	μ1 - μ2 < -|SM|
Superiority Margin (SM):	5
μ1:	112.8171
μ2:	126.3647
δ:	-13.54763
σ1:	17.3775277395911
σ2:	24.5118275915271

―――

				Efficacy		Futility	
Stage	N1	N2	t-Test Value	Boundary	Probability	Boundary	Probability
1	40.00	48.00	-1.9252	-5.3532		0.1854	
2	82.00	85.00	-2.6069	-3.2919		-0.7236	
3	*121.27	*121.27		-2.6606	0.8015	-1.2064	0.0005
4	*158.56	*158.56		-2.2908	0.1715	-1.6256	0.0010
5	*195.86	*195.86		-2.0422	0.0225	-2.0422	0.0055

―――

* Simulation sample size (Non-integer sample sizes were rounded to the next highest integer.)

Boundary Probabilities for δ = -6

Given the data that have already accumulated in the first two stages, and assuming means of 120 and 126 and standard deviations of 17.376 and 24.512 going forward, this report gives the simulation probabilities of crossing each of the future boundaries. The sum of the efficacy boundary probabilities is another estimate of the conditional power.

Boundary Probabilities for δ = -6
───

Number of Simulations:	2000
User-Entered Random Seed:	7970145
Futility Boundaries:	Non-Binding
After Efficacy Boundary Crossing:	Hold Out
After Non-Binding Futility Boundary Crossing:	Leave In
Alternative Hypothesis:	μ1 - μ2 < -|SM|
Superiority Margin (SM):	5
μ1:	120
μ2:	126
δ:	-6
σ1:	17.3775277395911
σ2:	24.5118275915271

				Efficacy		Futility	
			t-Test	───────────────		───────────────	
Stage	N1	N2	Value	Boundary	Probability	Boundary	Probability
1	40.00	48.00	-1.9252	-5.3532		0.1854	
2	82.00	85.00	-2.6069	-3.2919		-0.7236	
3	*121.27	*121.27		-2.6606	0.2245	-1.2064	0.0405
4	*158.56	*158.56		-2.2908	0.1815	-1.6256	0.2805
5	*195.86	*195.86		-2.0422	0.1185	-2.0422	0.5615

* Simulation sample size (Non-integer sample sizes were rounded to the next highest integer.)